DOCUMENT RESUME

ED 215 442 EA 014 589

AUTHOR Bernardin, H. John; And Others
TITLE The Applicability of Organizational Effectiveness Models for School Systems. Final Report.
INSTITUTION Virginia Polytechnic Inst. and State Univ., Blacksburg.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Jun 81
CONTRACT NIE-053-363933-1
NOTE 155p.

EDRS PRICE MF01/PC07 Plus Postage.
DESCRIPTORS *Educational Research; Elementary Secondary Education; Goal Orientation; Measurement Techniques; *Models; Organizational Effectiveness; Outcomes of Education; Research Design; Systems Approach
IDENTIFIERS *School Effectiveness

ABSTRACT

Researchers review the theoretical and empirical literature to investigate how models of organizational effectiveness can be applied to the study of school effectiveness in public elementary and secondary education. The authors consider nine models grouped into two sets: goal-centered models and systems-oriented models. The goal-centered group comprises the rational-goal, management-by-objectives, functional, and cost-benefit models. The systems-oriented concepts include the systems-resources, functional-structural, managerial process, organizational development, and Likert-ISR models. The authors discuss each model's applicability to education, problems in applying the model, and the outcome variables the model uses to measure school effectiveness. Four problems in the models are identified: (1) preoccupation with outcome instead of process variables, (2) use of deficient criteria of effectiveness, (3) lack of empirical support, and (4) lack of clarity regarding major variables. The authors propose a process-based model of effectiveness that includes variables related to organizational climate and human resources. After discussing past research designs for investigating school effectiveness, they suggest a research design that considers issues of data aggregation, sample size, and the validity of outcome criteria. (Author/RW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED215442

FINAL REPORT

The Applicability of Organizational Effectiveness Models for School Systems[1]

H. John Bernardin

Barry J. Riegelhaupt

Lea M. Curry

Virginia Polytechnic Institute and State University

Contract 053-363933-1

National Institute of Education

Washington, D.C.

Running Head: School Effectiveness

June, 1981

EA 014 589

Table of Contents

	Page
Abstract	1
Project Purpose	3
Organizational Effectiveness - An Introduction	7
Goal - Centered Models	10
Rational-Goal Model	13
Performance Contracting	16
Management-by-Objectives	24
Goal Measurement	36
Standardized Test Scores as Criteria	37
The Functional Model	48
Goal Attainment Scaling	50
The Cost-Benefit Model	53
Human Resource Accounting	61
Systems Models	65
The Systems Resource Model	67
The Functional-Structure Model	73
The Managerial Process Model	81
The Organizational Development Models	90
The Likert-ISR Model	95
Conclusion - The Choice of a Model for School Systems	98
Human Resource Process Variables in the Process Model	101
Research Designs in Studies for School Effectiveness	104
References	115
Tables	

The Applicability of Organizational Effectiveness Models for School Systems

Abstract

The purpose of this project was to consider the applicability of major models of organizational effectiveness for school systems and to consider research on school systems in the context of the models. Eight models were reviewed (i.e., rational-goal, management-by-objectives, functional, cost-benefit, systems-resource, functional-structures, managerial process, organizational development). While all models had features that were applicable to school systems, all also had major flaws as models to guide future research on school effectiveness. The most common problems characterizing the models were: 1) a preoccupation with outcome measurement with little consideration of "process" variables that lead to outcomes (e.g., management-by-objects, cost-benefit); 2) the use of contaminated and deficient criteria such as pupil gain scores on norm-referenced ability tests or measures of job satisfaction (e.g., functional, systems-resource, organizational development, Likert-ISR); 3) a lack of empirical support for the model (e.g., managerial process, functional-structures, functional); and 4) a lack of clarity regarding major variables related to the model (e.g., functional-structures, systems-resource). The management-by-

objectives and the Likert-ISR models were supported by some empirical evidence within school systems but all studies involved deficient criteria. A "process" model of effectiveness model is proposed which takes features from several models, both goal and systems-oriented. Organizational climate variables are included in the model along with components of the human resource selection, training, and evaluation processes for school systems. After critiquing research designs used in the past to investigate school effectiveness, a design is proposed that would be more sensitive to the most salient factors related to effectiveness.

The Applicability of Organizational Effectiveness Models for School Systems

The purpose of this report is to consider the applicability of the major models of organizational effectiveness for school systems. Specifically, the following models of effectiveness will be discussed: 1. rational-goal; 2. management-by-objectives; 3. functional; 4. cost-benefit; 5. systems-resource; 6. functional-structures; 7. managerial process; 8. organizational development; and 9. Likert-ISR.

When considering the appropriateness of the models for school systems, some difficulty was encountered in classifying particular educational research as representative of one particular model. For example, research on performance contracting will be discussed in the context of the rational-goal model of effectiveness though arguments could be made for its applicability to other goal-centered models such as management-by-objectives. The principal rason for this difficulty lies with the conceptual similarities of several of the models (e.g., functional and functional-structures; rational-goal and MBO; systems resource and cost-benefit).

Also considered in the discussion of models are the various outcome variables that are most often used to operationalize the construct of organizational effectiveness. The appropriateness of these operationalizations for educational systems will be discussed. It was the original intent of this project to evaluate criteria of school system effectiveness in order to arrive at a set of recommended strategies for data collection. However, a review of the literature in education revealed a paucity of detailed discussion of criteria other than student performance on ability or achievement

tests (e.g., Madaus, Airasian, & Kellaghan, 1980). There will then be only limited discussion of the criteria in terms of psychometric issues such as reliability, freedom from bias or contamination, and validity.

Variables from the educational literature that have been classified into one of the effectiveness models will be considered where sufficient data or discussion are available. For example, the use of student scores on achievement ability tests as results data in an MBO model will be discussed. Additionally, the appropriateness of the various units of analysis available to the researcher in education will be considered vis a vis the models of effectiveness.

Bidwell (1965) discussed the paucity of systematic research on schools as organizations. There has been an increase recently in empirical research on the functional, structural, and technological processes related to educational systems. Unfortunately, research on school effectiveness, like its counterpart in the literature on organizational effectiveness at large, has been largely fragmentary, discontinuous, and unidirectional. School effectiveness is almost always viewed in terms of one criterion and that most frequently is some measure of student outcomes. While there has been a plethora of position papers supporting programs of accountability, performance contracting, program evaluation, competency-based assessment, performance objectives, and organizational change, well controlled empirical research on models and/or the measurement of effectiveness is rare. This may be due to the fact that early large scale projects relating school characteristics to outcome measures bore little fruit. In one of the first major works in this area, Coleman, Campbell, Hobson, McPartland, Mood, Weinfeld, and York

(1966) concluded that "schools bring little influence to bear on a child's achievement that is independent of his background and social context" (1966, p. 325). Subsequent reanalysis of the Coleman et al. data and other similar studies have arrived at essentially the same conclusion (e.g., Jencks, Smith, Acland, Bane, Cohen, Gintis, Heyns, & Michelson, 1972; Mosteller & Moynihan, 1972; Postlethwaite, 1975; Smith, 1972). Research on educational interventions such as Project Talent, Follow Through, and Head Start is also in line with the statement above by Coleman et al. (1966) (e.g., Averch, Carroll, Donaldson, Kiesling, & Pincus, 1972; Jencks & Brown, 1975; Madaus et al., 1980). There are some studies that have found school characteristics -- accounting for significant portions of the variance in student outcomes (e.g., Bidwell & Karsarda, 1975; Cirirelli, Evans, & Schiller, 1970; Hanushek, 1977; Mayeske, Cohen, Wisler, Okada, Beaton, Proshek, Weinfeld, & Tabler, 1972; Mitchell, 1977). However, criticisms of many of these findings have been made on statistical or methodological grounds (e.g., Alwin, 1976; Cronbach, 1976; Datta, 1976, Hannan, 1971; Hannan, Freeman, & Meyer, 1976; Lau, 1977; Madow, 1969; Young & Erbring, 1976).

In summarizing the long list of point-counterpoint articles, Madaus et al. (1980) noted that "insights into the schooling process itself and the way in which that process is best viewed were often slow to emerge. Much of the debate has overlooked questions about schooling and school effectiveness and the most useful manner in which to conceptualize these concepts" (1980, p. 44). Hopefully, one or more of the models to be discussed on the following pages can serve as a framework for this conceptualization.

Before discussing the appropriateness of the various models of organi-

zational effectiveness for school systems, it is first necessary to put this research in proper focus regarding two major issues. The first issue has to do with the types of schools to be considered in the context of the models of effectiveness. There are great differences between the various institutions that come to mind when thinking of schools in general. The discussions to follow will focus on the appropriateness of the effectiveness models for public elementary and secondary mainstream school systems. The inclusion of private schools or colleges and universities would necessitate major critical diversions from the more generic characteristics of school systems. However, where empirical research seems particularly pertinent, some discussion of other educational systems will be made (e.g., the use of cost-benefit analysis for special educational programs; college student evaluation as criteria for teacher effectiveness). Also, an excellent discussion of organizational effectiveness for higher education can be found elsewhere (Cameron, 1978).

The second issue has to do with the comprehensiveness of the discussion of major topical areas such as accountability, performance contracting, criterion-referenced instruction, instructional or performance objectives, program evaluation, cost accounting, competency-based programming, teacher evaluation, educational assessment, differentiated staffing, PPBS, PERT, MIS, and a host of others. It was never the purpose of this project to review all of this literature that is obviously at least tangentially related to a ubiquitous construct such as organizational effectiveness. Rather the focus in the discussion to follow is on the applicability of the various models and criteria of effectiveness. Where research related to accountability, contracting, etc. fits within one of the models, it will be

included. An attempt has been made to incorporate most of the better research and discussion of such topics into the paper.

Discussions of the various models will differ substantially in length due either to the appropriateness/inappropriateness of the model for school systems and/or the extent to which research or discussion in educational literature indicates support/nonsupport for the model. Also, there may be longer discussions of non-educational research for some models when that research is considered pertinent and there is little educational research available (e.g., the managerial process model).

This report will begin with a brief introduction to the concept of "organizational effectiveness" and point out some of the difficulties with measuring it. The various models of effectiveness will then be clustered and discussed within either a goal-centered or "systems" perspective on effectiveness. To that end, the authors are indebted to the excellent work of Campbell, Bownas, Peterson, and Dunnette (1974) who summarized and critiqued the plethora of research on effectiveness up through 1973. Research and discussion subsequent to this date will be the focus of this report.

Organizational Effectiveness - An Introduction

There is a notable lack of consensus as to what the construct of organizational effectiveness actually means (Steers, 1976). Pennings (1976) notes that the literature is rather diverse in its prevailing definitions of organizational effectiveness. The research and debate is, for the most part, noncumulative and disjointed. There is agreement that organizational effectiveness is a desirable state of affairs, but there is concomitantly a lack of firm theoretical basis for measurement of the construct. Ghorpade

(1971) states that a researcher undertaking a study of organizational effectiveness is faced with two major hurdles at the theoretical level: (a) there does not seem to be a substantive, systematic body of knowledge based on hard evidence about the functioning of organizations; and (b) there does not seem to be a universally accepted theoretical framework for approaching the study of organizations. Unfortunately, these hurdles also apply to research in education. If organizational theory is itself still in such a confused state of affairs, it is small wonder that studies of the effectiveness of these structures are in such turmoil.

This state of turmoil is clearly illustrated by the multiplicity of criteria that have been used to assess organizational effectiveness. A large and often conflicting array of data has been amassed and purported to demonstrate effectiveness. Campbell et al. (1974) note 26 separate criteria that have been used at one time or another in the literature. They conclude that a better way to think of organizational effectiveness is as an underlying construct which has no direct operational definition, but which constitutes a model or theory of what organizational effectiveness is. The function of such a model would be to identify the variables of interest and tell us how they are related.

In part due to the lack of agreement concerning the criteria of organizational effectiveness, it is no surprise that there is also conflict concerning the best methods for its attainment. Steers (1976) notes that a major reason for this lack of agreement is the parochial view that many researchers take toward the construct. Effectiveness is often viewed in terms of a single criterion that is person or situation specific (e.g., profit,

efficiency, output, etc.), even though organizations are known typically to pursue multiple goals that are modified across personnel, technologies, socio-cultural environments, and a host of other factors. Yet, most often, the attainment or failure to attain one goal is treated as a dependent variable and the larger context is ignored. As will be discussed below, research on school systems' effectiveness can also, for the most part, be characterized as parochial in nature.

A second reason for the lack of agreement concerning effectiveness attainment is the ambiguity of the concept itself. Many models of organizational effectiveness are somewhat esoteric and the criteria are difficult to operationalize. Both the models and the criteria may be person-viewpoint specific, dependent mainly on the frame of reference of the evaluator. Steers (1977), for example, reviewed 17 different studies encompassing as many different multivariate models of organizational effectiveness. He found that only one criterion (adaptability-flexibility) was mentioned in more than half of the models. Thus, there is little agreement about the conceptualization or measurement of organizational effectiveness. In education, there is a general consensus that some measure of student outcome is the most appropriate criterion on which to assess effectiveness. However, there is also a consensus that all such measures have problems (Madaus et al., 1980). The literature on organizational effectiveness is most often divided theoretically into two major perspectives. While the nomenclature varies somewhat, the two most popular labels for those perspectives are the goal-centered and the systems view (Campbell et al., 1974; Ghorpade, 1971).

The framework for the discussion of the specific models within these

two perspectives is as follows: 1) a general discussion of the meaning of the model, and definitions of all pertinent terms; 2) a consideration of the applicability of the model for school systems; 3) a review of pertinent research in education; and 4) (where possible), a discussion of implications of the model for the practitioner in education.

Goal-centered Models

The goal-centered models make the assumption that the organization is in the hands of a set of rational decision makers who have in mind a set of goals that are reasonably explicit and for which the organization mobilizes its resources (Campbell et al., 1974). It is further assumed that the goals are reasonably well defined and few enough in number to be manageable. Given this, it should be possible to develop a strategy for the attainment of maximization of these goals. Effectiveness criteria would then be defined in terms of the degree to which stated goals were achieved, or as Price (1972) views it, the degree to which the organization achieves ideal end-states.

Steers (1976) notes two basic variations of goal-centered theories that dichotomize the field. The model proposed by Thompson and McEwen (1958) incorporates those factors external to the organization that can influence the selection and modification of organizational objectives. This model focuses on the relationship between an organization and its environment in reference to the balance of power. The second model, by Cyert and March (1963), deals principally with internal forces. It focuses on the internal determinants of goals by coalitions of individuals and groups. This dichotomy of focus between internal and external holds true for many of the models

to follow. Some models, however, (e.g., MBO, cost-benefit) are influenced by both internal and external forces in the selection and modification of objectives.

The goal-centered approach views organizations primarily as rational devices for the attainment of certain objectives. Goals answer the question of where the organization is going. Steers (1976), however, questions this simple conceptualization of goals on several logical and methodological grounds. For example, one can ask questions such as whose goals are we talking about? In the context of school systems, is the reference group teachers, students, parents, administrators, or whom? There is also the problem of multiple and often conflicting goals. Who decides which goals are more important, and how is this decision made? Are they system-wide goals or strictly school-specific goals? What is the time reference? Given limited resources, what will the organization do? Thus, according to Steers (1976), the simple definitions of goals that are usually given may be useful on a fairly abstract level, but they lack the precision necessary for examining empirically the relationship between organizational goals and organizational effectiveness. Yuchtman and Seashore (1967) agree, stating that goals as ideal end states do not offer the possibility of realistic assessment.

At this point a distinction must be made between the various types of goals toward which an organization may be striving. The most common distinction is that of official, operative, and operational goals (Perrow, 1961). Official goals are formal statements of purpose made by executive management concerning the nature of an organization's mission. These goals are typically vague, aspirational, and set in an infinite timetable.

Operative goals represent the real intentions of an organization; that is, what an organization is <u>really</u> trying to accomplish regardless of what it says it is trying to do. Operational goals are goals for which there are agreed upon criteria for evaluating the degree to which an organization's activities are contributing to goal attainment. Goal attainment here is operationally defined in terms of unit production, profit levels, etc. Steers (1976) and Reimann (1975) both agree that assessment of organizational effectiveness must focus on operative goals. If it does not, then the concept of a goal-directed system has little meaning.

One shortcoming of the goal approach is that some goals arise as cultural entities, derived from societies' expectations and demands, and thus are not applicable for analysis. Of particular importance here is the concept of goal relevancy. Although goals may be important to some segments of the organization (and, by implication, to some segments of society) they may not be to others. In fact, some segments of the organization may attain their goals at the expense of other segments. Seldom, however, do we examine this multi-goal environment within which the organization operates (Kirchhoff, 1977).

A second shortcoming is that the approach has limited usefulness for studies of comparative effectiveness across organizations. If we accept the notion that the only valid indicators of an organization's effectiveness are its operative goals, then measures of effectiveness will be different for each organization. Although this approach may have practical utility for any given case, it would be prohibitively restricting for any theory of organizational effectiveness.

Indeed, Price (1972) summarizes his review of goal approaches by concluding that they have not successfully measured organizational effectiveness. Kirchhoff (1977) notes further that current empirical methodologies do not adequately express the complexity of effectiveness measurement.

A third issue concerns the actual measurement of goals (i.e., the "criterion problem"). This issue will be examined in depth in the management-by-objectives section.

Presented in Table 1 are four goal-oriented models of organizational effectiveness. We now turn to a discussion of the applicability of each of those models as a vehicle by which to assess a school system's effectiveness.

Insert Table 1 about here

The Rational-Goal Model

The basis of the rational-goal approach is the Weberian concept of functional rationality. According to Weber (see Steers, 1976), modern organizations are characterized by networks of roles, divisions of labor, and hierarchies of activities (i.e., a typical bureaucracy). When an organization is effective, every element has a defined role or function that is logically related. The organization is evaluated by comparing those goals that were established with those that were actually accomplished. Thus, an effective organization is one that has achieved the goals that it has set for itself (e.g., Etzioni, 1964; Perrow, 1961). Criteria are established based on either the formal or informal (operative) goals of the organization (Cunningham, 1977).

In a discussion of elementary education, Lortie (1969) maintains that goal statements are generally stated at the level of the school board and due to their general nature, it is difficult to define criteria which would determine the extent to which the goals were successfully achieved. Dornbusch and Scott (1975), however, argue that the goals at the elementary level could be far more explicit than goals established, for example, at institutions of higher education. Cameron (1978), on the other hand, shows empirically that "dominant coalitions" from six institutions of higher education can agree for the most part on a list of important criteria on which to assess effectiveness. It may very well be that disagreement, however, would flourish as soon as any detail was required for the ambiguous criteria reported by Cameron (1978) such as "student educational satisfaction", "academic development", or "professional development and quality of the faculty".

Numerous criticisms have been levied against the rational-goal model of effectiveness. Blau and Scott (1962) and Scriven (1967) state that there is a preoccupation with formal or official goals to the exclusion of multiple constituencies affecting organizations with their informal, operative goals. While this criticism may have been true of goal-oriented, educational models in the past, there has been a substantial increase in consideration of operative, and more quantitative goals of school systems from relevant constituencies (e.g., Nash, 1970). Thus, whereas goals may have been defined earlier as, for example, "to impart skills in languages and mathematics" (cf. Beshers, 1972), recently the method of assessing performances in these areas and the extent of improvement are very often made explicit (e.g., Barro, 1970). The increased interest in performance contracting is a

reflection of this greater concern for actual performance measurement to assess goal achievement. One of the better examples of an educational study to assess operative goals is the large scale project on performance contracting sponsored by the Office of Economic Opportunity in 1970-71. Due to its implications in assessing the rational-goal model of effectiveness, discussion of this study will follow shortly.

A second common criticism of the rational-goal model is that there is a general neglect of the conflict or potential conflict of many organizational goals (e.g., Rice, 1963). This criticism is quite appropriate for school systems given the multiple constituencies concerned with their effectiveness (e.g., politicians, community leaders, teachers, parents, students). The goals of these various constituencies will certainly conflict at times, as, for example, in considerations of teacher salary and absenteeism policy, budget factors, or extracurricular activities.

A third criticism of the rational-goal theory is that the interactions of an organization's goals with the goals of the affected environment are not taken into consideration (Lawrence & Lorsch, 1969). Thus, for example, operative goals of the school system such as "equal employment opportunity" may be influenced by the informal goals of the organization's boundary institutions such as the local or regional affirmative action office or the racial make-up of the school board. Also, environmental factors such as the socio-economic characteristics of the school community can have a significant impact on the types of goals and the measurement strategies established to assess goal attainment (Averch et al., 1972).

A final criticism of the model is that it treats goals as static in an

environment that realistically should make all operative goals dynamic (Pfeffer, 1977). Although the formal goals of a school system may be constants, the operative goals may be affected by contextual factors. For example, the operative goals of teachers could change significantly as a function of some budgetary decision (e.g., working to contract due to low or no pay raises).

Performance Contracting. With performance contracting, the focus is on the school systems', teachers' or outside agencies ability to achieve certain explicit goals related to educational outcomes (Gramlich & Koshel, 1975). Thus, goals are formalized in terms of end states, a time reference, and "hard" criteria (e.g., average scores at or above a certain score on a specific test or tests). For example, in 1969, Dorsett Educational System contracted with two school systems (in Texarkana, Arkansas and Liberty-Eylau, Texas) to teach certain academically deficient high school students. The performance contract explicitly stated that Dorsett would be paid only for those students who attained a specific level of proficiency. The program involved more than 200 students, and seemed to work amazingly well, exciting many in educational circles to the possibilities of performance contracting as a way of making schools more accountable for the outcome of their students. However, the validity of test score gains in Texarkana has since been questioned. Dorsett has been charged with contamination of the results by teaching specific items on the test to students (Educational Research Service, 1974).

As a result of the Dorsett study, the Office of Economic Opportunity (OEO) decided to perform a national experiment designed to assess the

relative advantages and disadvantages, as well as the practicality of performance contracting for use with academically deficient children.

The evaluation instrument used was the Metropolitan Achievement Test, a battery of tests including reading and mathematical components.

Of an original sample of 24,000 students, 81 percent actually participated in the entire program with only 61 percent having scores which could be evaluated. There were slight differences between the number of students in the control and experimental group who were both full time students and took pre- and post-tests (74 percent of the control groups as compared to 77 percent of the experimental group). Gramlich and Koshel (1975) stated that these differences "suggest that a disproportionate attrition could be responsible for a small bias against the students in the experiment, especially if the experiment was itself responsible for the lower attrition in that group" (p. 32).

The differences between the various subtests in both experimental and control groups were relatively small, showing that the students did not demonstrate a tendency to do better in one area than any other. One interesting difference found was that in the subjects not taught by the contractors (spelling, word analysis, etc.) the experimental students did slightly worse. One reason for these differences might have been that students in the experimental group may have missed some instruction in other subjects due to ineffective and haphazard scheduling or other disruptions. Other interpretations, of course, are possible. For example, teacher bias may have affected the results. The teachers could possibly have been biased against those students in the experimental group, thereby giving them less attention.

This could easily be the case, for the idea of performance contracting was viewed as a threat by many teachers. As will be discussed later, many of the teachers were indeed uncooperative and sometimes hostile in regard to the program.

Since in most cases achievement gains were positively related to the pretest scores, and since these scores were not equivalent in experimental and control groups (the experimental group was usually lower) these gains were adjusted. The adjusted mean gains are perhaps the best measures of the actual differences between the two groups. "They show statistically significant experimental gains four times, one statistically significant experimental loss, and very small gains or losses in the other seven cases" (Gramlich & Koshel, 1975, p. 39).

Overall, the adjusted experimental gains average about 0.04 grade equivalent units, a seven percent gain over that of students in the control group. These results were disappointing to those involved, and forced OEO to conclude that performance contracting proved no more effective than regular teaching methods in improving the level of academically deficient students.

Although OEO's findings did not support performance contracting, many have pointed to numerous flaws in the design, procedure, and implementation of the experiment. The first of these flaws had to do with the lack of time spent organizing the experiment.

Also criticized was the length of the experiment. It has been suggested that one year is not enough time to evaluate the amount of change performance contracting is capable of produing. Indeed, according to some, this lack of

extensive evaluation prevents the results from demonstrating any conclusive evidence whatsoever. Thus, "the OEO experiment showed that private enterprise could not provide existing schools a royal road to educational quality within one year . . . Beyond this conclusion, however, the experiment left many questions unanswered" (Carpenter-Huffman et al., 1974, p. 52).

A third time-related problem deals with differential amounts of time devoted to experimental and control groups. It would seem then that the differences in mean achievement gains between experimental and control groups did not depict a very accurate account of the relative merits of performance contracting. The experiment would have been much more informative had this time element been held constant.

A more serious criticism of the OEO experiment has to do with the procedure for designating schools as "experimental" or "control." As previously stated, this status was determined on the basis of academic need. Consequantly, the control and experimental groups were not evenly matched in terms of ability. This uneven matching constitutes a serious flaw in the experimental design. In particular, a design with nonequivalent groups runs the risk of maturation effects being mistaken for the effect of the variable, thus affecting the internal validity of the experiment. The external validity is also threatened by the effect of the interaction of testing and the treatment. If either of the groups are selected for their extreme scores, the difference in the degree of shift from pre- to post-test may be a product of regression rather than the effect of the experimental condition. Indeed, in their statistical analysis of the experiment, Carfinkel and Gramlich caution that "one must be extremely careful in interpreting experimental

results if there is some indication that the sample is imperfectly matched" (Garfinkel & Gramlich, 1973, p. 284).

Another area of great concern with respect to the experiment involves perhaps the major problem with rational-goal programs such as contracting, namely, that of the criteria that are used. The contracted firms questioned the content validity of the criteria, claiming that the Metropolitan Achievement Test did not sufficiently match the content of their program. Indeed, in addition to the firms, many have questioned the use of standardized achievement tests to assess present capabilities of a student (see for example, Brophy, 1973; Hall, 1964; Soar & Soar, 1977). While some of this discussion will be reserved for the MBO section of the paper, a comment must be offered on the suitability of standardized, norm-referenced tests as criteria in contracting and other accountability programs. The validity, reliability, and relevance of such tests for use as criteria of school or teacher effectiveness has been questioned by numerous writers (e.g., Airasian & Madaus, 1976; Coleman, 1975; Kellaghan, 1977; Lewy, 1973; Madaus, Kellaghan, Rakow, & King, 1979; McClelland, 1973). Madaus et al. (1979) ask the rhetorical question whether verbal ability tests such as those frequently used in contracting or voucher programs adequately sample the content domain that schools attempt to teach. Cooley and Lohnes (1976) contend that while such tests are certainly not perfect, they are the best thing we have available. Others (e.g., Popham, 1970) recommend a form of criterion-referenced testing to assess school or teacher effects. Madaus et al. (1979) agree, stating that tests should be used that are more congruent with the specific school curriculum and instructional methods.

Additional problems with the OEO study concern the physical conditions under which students were tested at certain sites, and the participation (or lack thereof) of parents, students, and in particular teachers. Many teachers opposed the program because they felt threatened by the entire performance contracting idea. It seems quite reasonable to suspect that some of this dissatisfaction with the project may have biased teachers against students in the experimental program. If this were so, one might expect to find that these students did not perform well on other subjects taught by the teacher. In fact, this result was found.

Given the lack of planning time, unsatisfactory methods of measurement, insufficient assessment time, and poor research design, it seems the only conclusion one can draw from the OEO study is that there was a lack of planning time, unsatisfactory methods of measurement. Thus, one can say that this experiment did not support the idea of performance contracting, which is not to say that "performance contracting must therefore not work."

The OEO and Dorsett studies underscore some of the problems with the rational-goal model of organizational effectiveness as applied to school systems. While operative, quantitative goals are stated, they typically do not represent the domain of formal goals for a school system. The result, as in the OEO study, is a preoccupation with those goals which can be formally measured. The conflicts of multiple constituencies regarding goals are also evident in the performance contracting framework. A concentration on certain educational outcomes as end-states would inevitably result in conflicts over appropriate means for facilitating such end-states between students, parents, principals, and teachers. The hostility of teachers to

the notion of performance contracting is also typical of reactions to goal-oriented management approaches which do not base specific end states on input from those most affected (Dornbusch & Scott, 1975).

The discussion of the OEO project was presented to accentuate some of the problems common to performance contracting in general and to a rational-goal approach to educational accountability. Similar critiques could be made of other rational-goal approaches to accountability programs such as voucher plans (e.g., Fobert, 1977). Critics of contracting view such problems as either insurmountable (e.g., Hottleman, 1971) or in dire need of attention before wholesale adoption of such an approach (e.g., Gramlich & Koshel, 1975; Mayrhofer, 1972). Mecklenburger and Wilson (1971) present a case study of some of the problems with performance contracting in a large city system. As in the discussion of common criticisms of the rational-goal model of effectiveness (e.g., Blau & Scott, 1961; Cameron, 1978; Rice, 1963), Mecklenburger and Wilson (1971) cite the conflict of operative goals across the various educational constituencies as the principal cause of discord with contracting. Teachers tend to view performance contracting as a punitive measure to be used only as a basis for firing people (Gramlich & Koshel, 1975). The use of contracting in any sort of teacher incentive framework is rare and not well received by teachers anyway (Nation's Schools, 1970). The reaction is not unlike the common reactions to other goal-oriented programs such as management-by-objectives (Levinson, 1970).

A major criticism of the rational-goal model that appears to be adequately handled by school accountability programs such as contracting or vouchers is the belief that organizational goals are retrospective in nature

and serve only to justify action, not to direct it (Weick, 1976). Under most accountability systems there are at least formal, long-term goals set with changes made in criteria for effectiveness prior to teaching. For example, McDonald's (1972) accountability system for New York City included a provision for periodic review and revisions of all program objectives (see also Alkin, 1973; Barro, 1970; Lieberman, 1970; and Lopez, 1970). Thus, many accountability programs do view schools as open systems where goals and criteria are changeable.

For the practitioner concerned with assessing effectiveness with the goal of improvement, the problems of conflicting goals and constituency involvement are person-specific, not approach-specific, and can be adequately handled by a "cascading process" in goal-setting. In this approach broad objectives are set at the school board level and filtered down into specific goals at the school level. In this way all who are to be affected take part in the goal setting process. This should result in school-level goals that are compatible with broad system objectives and at the same time foster a feeling of cohesiveness throughout the entire school system because all involved are working toward the same goals.

There remain, however, some problems that are approach-specific making the rational-goal approach questionable as a framework within which to assess a school system's effectiveness. First, the approach is concerned only with outcomes. There is no mention of <u>how</u> goals are to be accomplished, only that they should be. This outcome orientation results in a second problem. The rational-goal approach, by concentrating on "ends", will by its very nature, emphasize those skills that are more easily measurable, at the

exclusion of other skills. As a practical matter, this may result in a concentration by teachers on the basic skills which are measurable by standardized achievement tests at the expense of more abstract skills. This brings us to the final problem - the measurement issue. This issue along with the two previously mentioned criticisms will be returned to after a description of the management-by-objectives approach since they are relevant to both approaches.

Management-by-Objectives Model

Many researchers view management-by-objectives (MBO) as a complete system of planning and control interfaced into the organization (see, for example, Campbell et al., 1974 and Odiorne, 1965). However, the basic program of MBO originated as a mechanism, or device, the activation of which would result in increased performance and organizational effectiveness by carefully formulating organizational goals in terms of quantifiable results and the behavioral directives needed to achieve them. This initial formulation, though, was soon augmented by a need to confirm results, and to implement cooperative planning, and behavioral intervention, as initial attempts to engage the "mechanism", were seen to fail. The concept of an MBO system replaced the original mechanistic formulation, and some researchers today go so far as to say that MBO is a complete philosophy of management (Albrecht, 1978).

Peter Drucker, in The Practice of Management (1954), proposed a principle of management for business enterprise to harmonize the goals of the organization while allowing for full individual output. This was presented in terms of results accountability as opposed to the emphasis on the human

relations approach dominant at the time. However, although Drucker had implied a philosophy of management, the rigid or careless applicability of this principle without regard to human resource factors, or incomplete implementation in terms of evaluation deficits, demonstrated the limitations of a mechanistic MBO. Numerous criticisms and new suggestions have been offered. A system of MBO was deemed necessary -- a way to incorporate MBO into the conceptual framework of organizations as open systems of the total system, so that the objectives are a coordination of the goals of all the other subsystems in light of the organization's main goals. Since the task is to quantify these objectives, these results could be considered as dependent variables (or functions) of the parameters (systems) of the organization. Of course, the problem lies in the delineation of parameters for the more tenuous constructs, such as job satisfaction and the need to recognize human relations factors in any MBO formulation. In theory, however, an MBO system should maximize the efficiency of the organization by not only the increased incremental sum of the achieved goals, but by the interaction of the entire system as well, creating a set of proven methods which together make a total and comprehensive management system to improve company and manager performance (Humble, 1973). System approaches toward the implementation of MBO have increased their attention to parameters such as job satisfaction and other human relations factors in the recognition of the need for a cooperative thrust at the subordinate employee level. Research in this area attempts to quantify these more abstract constructs with more measurable effects. Systems approaches are the most prevalent interpreation in the recent literature, in which the MBO system functions as an organic

tool to be properly interfaced: "when the various aspects are integrated with a balance between the individual and organization needs, ... you have organizational effectiveness and results" (Beck, 1976).

Those who view MBO as a philosophy tend to subsume the previous approaches in the formulation of an MBO attitude to be conveyed to subordinates by the manager in a holistic way. Karl Albrecht describes this philosophy in his Successful Management by Objectives (Albrecht, 1978) as a basic mentality of managing and that when mistaken for a system or a set of principles, a "rigor mortis" mentality can develop which will destroy the adaptive features of efficient management by allowing the responsibility of management to slip into the "mechanism" or system. In other words, Albrecht believes that when MBO is taken to be a "system" by which to manage effectively, the possibility of inappropriate initial installment and failure to accommodate changes in the organizational environment can obscure the effectiveness of management. Critical to Albrecht's formulation, the MBO philosophy consists of: (1) an objective-oriented manager whose "systems" understanding conforms to the organizational goals and guides subordinates; (2) the need to instill the appropriate objectives orientation to the workers (with a full understanding of what MBO is trying to achieve); and (3) the implementation of the above principles in a reward-centered environment.

MBO, like the rational-goal model, specifies the primary criteria of effectiveness by checking whether or not the organization has accomplished its specific goals, and thus fits closely under the goal-oriented operationalization of organizational effectiveness. MBO, as stated by Campbell et al. (1974), rather than evaluating the organization on some single abstract

continuum such as the cost/benefit ratio, states that effectiveness is some aggregation of specific, concrete, observable and quantifiable accomplishments or failures. Therefore, organizational effectiveness can be approached through a review of MBO studies on performance evaluation, and in terms of school evaluation, by focusing on MBO in the public sector, and where MBO or facsimiles already have been established in school systems (e.g., Dyer, 1970; Gray & Burns, 1979; Herman, 1978; Lasher, 1979; Malsam, 1979).

Management-by-Objectives implies a planning function as represented by the need to quantify what must be done (after careful analysis of why it must be done, from the organization's long-term and short-term goals), the specification of the necessary degree of detail as to how it must be done, the coordination of activities embraced by the process mandates assigning the time these tasks are to be done, and finally, the need for accountability suggesting a weighing of costs of the system versus the benefits.

Management-by-Objectives also implies a controlling function in the sense that it provides a method for determining how well the prepared objectives are being performed, how effective the completed activities are in realizing the organizational goals, and the necessary corrective action for improving performance or dealing with unexpected situations.

Jones (1977) considered the problem of accountability in education through an MBO procedure for guidance and counseling programs. The five month study involved 32 counselors and 572 students in a northwestern Chicago high school district including five high schools. An MBO technique was employed for the test group of counselors, and a traditional group and control group (whose members believed they had an "instinctive" flair for

answering student needs) were included. Reactive data indicated support for MBO but no "hard" criteria were used in the assessment.

A similar study of guidance programs in Maine by Vanzandt (1977) incorporated the use of a special scale developed to quantitatively measure the quality of the program's services. The scale was designed for three student groups and a teacher evaluation group by means of a Likert 5-point scale rating both the programs and services, and the counselors' personal/professional skills. The study was an intervention design where an MBO program was implemented after a baseline had been established, and feedback concerning the "treatment" during the course of the study was a critical element of Vanzandt's MBO model.

Random samples of students from the sixth, eight, eleventh, and twelfth grades and all teachers were included in the administration of the Guidance Service Rating Scale and a needs assessment scale to provide the information about each of four school's guidance needs. Steering committees composed of students, teachers, administrators, and counselors used the information to determine the priority concerns of guidance, then established criterion-referenced, behavioral objectives to direct the guidance programs toward the priority needs. Time phasing and strategies focusing on counselor-outcomes were incorporated to emphasize the intervention strategy and constrain the interpretation of the results. Communication about evaluation results and decision-making were stressed as integral elements of the MBO model. Unfortunately, due to the limited time that could be devoted to the project, the results were more a reflection of the process of using the MBO model than an evaluation of the effectiveness of MBO.

Although the problems with this design are the consideration of extraneous variables and the reliability/validity of the rating scale, the scale did find significant differences ($p < .05$) in terms of the evaluations made by the students and teachers from the Guidance Service Rating Scale. The author concluded that MBO may have a place in these programs for its practicality in helping guidance programs establish accountability.

In a study involving university faculty, Shetty and Carlisle (1974a, 1974b) found that their MBO program increased the awareness and specificity of goals for participants, improved career planning and professional development, and led to overall improvements in planning. Additionally, supervisors in the study indicated that the MBO program led to increased objectivity in performance appraisal. Perhaps most important was the finding that MBO led to "significant" increases in individual performance. On the negative side, participants indicated a substantial increase in "unnecessary paperwork." Unfortunately, the authors measured the effectiveness of the program with a post-implementation measure only, used no control group, and measured all variables with rating scales. Similar results were obtained by Smith (1975) in work with staff members of three universities. Two other studies (Lasher, 1979; Marsh, 1979) report positive results with MBO in academic settings but present no data to support their positions. In a study involving junior college administrators, Rossano (1975) found better supervisor/subordinate relations subsequent to the implementation of MBO but no difference in performance levels.

Applications of MBO at the elementary or secondary education level have been quite limited. Altergott (1970) collected interview data for high school

teachers involved in an MBO program. Participants indicated that MBO encouraged more specific program planning, improved subordinate/superior relations, increased subordinate participation in goal setting and perceived subordinate authority. Finally, participants indicated a higher level of motivation under the MBO programs. Altergott (1970) concluded that MBO is an excellent approach to management for education. However, his study was methodologically weak in that no control groups were used, no statistical tests were made, no objective data were used as criteria, and only post-implementation interviews were conducted. In a study to assess attitudes of participants toward MBO for appraisal purposes, Eads (1974) found elementary and secondary teachers expressed a more positive attitude toward a "supervision by objectives" program for evaluation than several other more subjective approaches. Principals also indicated a stronger preference for the MBO approach to appraisal. The study was one of the more methodologically rigorous designed to assess the effects of MBO. A longitudinal design was used with pre-test and post-test data and results were compared to results from a comparison group.

Herman (1978) and Malsam (1979) present case studies of MBO implementations in school districts. While no data were reported, Herman (1978) stated that objectives were successfully met with the program. Several of the objectives were concerned with cognitive achievement factors, such as: to decrease by a minimum of 3 percent the percentage of students reading below grade level in the middle school, grades 6 through 8; and to cause a gain of at least one month beyond the normal reading pattern established by the pupils who remain below grade level within the time period September 1, 1977-

June 30, 1978 (p. 89).

McLagan (1980) in studying a special education center found benefits in the MBO approach to defining measurable goals and specific, identifiable behaviors.

Alpin and Schoderbeck (1976) presented their evaluation of MBO in the public sector by listing the constraints found in this setting and offering three suggestions critical in the implementation of MBO. The suggestions were: (1) to begin with the initial formulation of clear strategic and operation objectives; (2) to determine the appropriate evaluational criteria; and (3) to establish and develop an information gathering capability.

The researchers say that while in private organizations the formulation of strategic goals is more of degree than a question of kind, in public organizations the strategic goals have never been stated adequately. They cite the results of the Senate Committee on Governmental Operations in describing seven major program areas and an administrative support area reflecting the need to define "key-result" areas for successful MBO. To fulfill these strategic objectives, another tactical level is required, which includes the short-run activities. At this point, the planners must develop a congruency between these levels.

Once this has been achieved, the criteria to evaluate the execution of these goals must be developed, and while goal attainment should be concerned with costs, quality is an important consideration as well. The researchers also related a number of requirements for criteria they felt necessary (e.g., definition of target, area of coverage, clientele penetration levels, economic restraints, time dimensions, estimation of major consequences, and a focus on

the specification of the types and quality of services). Unfortunately, the authors did not specifically state any suggestions for the quantifiability of these factors, but merely tried to identify the range and scope of these items of consideration.

Aplin and Schoderbeck collected information from public administrators employed by a state agency having over 7000 employees, and found a number of factors critical for successful MBO application within a public organization: 1) influence in goal setting; 2) relevancy of feedback; 3) top level goal setting; and 4) organizational support of MBO.

The study suggests that the first consideration, the goal-setting procedures, proved most effective when the subordinates contributed equally as much planning as their superiors, and least effective when they perceived themselves to contribute more planning than their superiors. The data taken from the surveys showed 53% of the respondents depicting this situation with only 19% reporting the optimum situation. The researchers suggest that in the public sector, there may be a tendency for a "too much" degree of freedom in establishing objectives -- a tendency for an overly liberal degree of freedom in allowing subordinates to establish objectives which will negatively influence these subordinate's perception of their roles.

Relevancy of feedback was another factor seen as critical for successful MBO. Although much literature on feedback had already been collected, the researchers point out that relevancy of feedback will induce more positive attitudes, greater managerial efficiency, enhance the goal-setting process, and, in general, generate more MBO support and effectiveness.

Another factor considered by Aplin and Schoderbeck was top level goal-

setting. They noticed a previous absence in the formulation of objectives at the top administrative level, and impairment to the resulting MBO program in these particular organizations.

The fourth consideration in successful MBO implementation was the organizational support of the program. The researchers point out that the interest of the administration must carry over to their subordinates, and specifically, the survey shows that the subordinate managers mirror the interests of their superiors.

Aplin and Schoderbeck also delineate some constraints on the use of MBO that have special significance for educational systems. First, they state there is reluctance on the part of public administrators to rely on objective data for evaluative purposes. A similar skepticism has been documented for principals when considering the use of achievement test data to evaluate teachers (Austin, 1979; Davis & Gross-Davis, 1974). Such a reluctance stands as a barrier to the widespread acceptance to MBO among educators. Principals, for example, may feel that they have acquired a special capacity to subjectively evaluate their teachers, and that an objective procedure such as MBO will interfere with what they consider to be their uniquely appropriate appraisal method. Another factor that resists the full implementation of MBO is the reluctance of administration to reward differential levels of teaching performance. The use of merit pay of any kind in elementary and secondary education has been rare with results equivocal (e.g., ETS, 1971). Additionally, educational systems are often controlled by political configurations that may undermine the actual control exercised by principals. Such a situation will, in turn, induce a reduced obligation toward compliance with

the MBO strategy. Finally, Aplin and Schoderbeck (1976) point out that today's changing environment demands more flexibility and innovation in the public sector than the private sector to deal with the changes in morality, politics, legislation, and technology. It is difficult to discern how such flexibility would affect objective criteria such as scholastic achievement and thus, MBO programs established with goals related to achievement.

Luthan (1976) has suggested that MBO shortcomings in the public sector are a function of the misapplication of the system with the different demands of service organizations. He proposes a contingency approach to management where functional if-then relationships between the environment (both internal and external) and management concepts and techniques that lead to effective goal attainment are determined. He maintains that MBO, as any other management concept or technique, cannot be universally applied to all situations, because of environmental differences. The identification of these environmental variables is thus necessary in order to implement MBO in the appropriate fashion. Luthan then suggests examining the popular constraints as a starting point, including: 1) the preponderance of pre-set rules, laws, and regulations; 2) less opportunities to participate in setting objectives; 3) the emphasis on seniority versus merit; and 4) the lack of adequate performance measures. He states that these factors are more often used as excuses for poor management, rather than impassible barriers to effective management.

While Luthan feels that the most urgent task is the identification of these variables, he presents no solutions or suggestions in light of his contingency management approach. Luthan says this type of examination must cover such factors as the nature of objectives and political (educational)

environments, high levels of turnover for public organizations, and the rapid change in policies inherent in these organizations.

Before an organization implements an MBO system there are some basic questions which must be answered: 1) what groups or individuals set the goals for a particular organization; 2) to what extent is it realistically possible to define quantifiable goals for an organization; 3) how should the relative importance of each goal be judged; 4) to what extent is it possible to know whether an objective has been accomplished; and 5) is the organization willing to commit the necessary time and effort to the MBO procedure?

There appears, however, to be an even more basic issue to be raised in considering MBO as a model of organizational effectiveness. Odiorne (1965) views the system of management-by-objectives as a process whereby the superior and subordinate (or in this case, the administrator and principal or the principal and teacher) jointly identify goals, define individual major areas of responsibility in terms of results expected of him/her, and use these measures as guides for operating the unit and assessing the contribution of each of its members. This definition, according to Albanese (1975), brings out two ideas basic to most applications of MBO. First, MBO implies an interaction between managers and their subordinates in goal setting. A second basic idea of MBO is appraisal of actual performance against the agreed-upon goals in the goal setting phase.

There is an argument that the objective of MBO is to help the individual teacher be successful because as the teacher succeeds, so does the principal, and the school, and so on. While the organizational literature cites the

benefits of goal setting for individual performance (Latham & Yukl, 1975), successful completion of individual teacher goals or objectives, however, do not necessarily guarantee a successful school system. To conclude that if a given teacher improves reading levels at the third grade by two months, automatically makes the school system as a whole effective, is to imply that individual teacher goals can be chosen so as to totally represent the goals of the entire school system. This is certainly not the case. While an individual teacher's goals can be made compatible with the system's goals they are not a substitute for them. All teachers in a school system may have attained their goals under an MBO system, but the school system itself may not have reached its goals. For example, many teachers may have set and reached goals concerning such aspects of the job as planning, managing, and organizing instruction, providing a favorable psychological environment for learning, and increasing reading levels by two months. However, if the school system that year ran at a substantial financial deficit, it would be difficult to conclude that the school system was effective even though certain individual teachers were effective. The point here is that the teachers' goals and the systems' goals may not be completely compatible. Most of the applications of MBO to school systems have made little attempt to adopt any type of "cascading process" in goal setting, that would foster a closer compatibility between goals across organizational levels. Most of the research has focused on teacher or principal goal attainment, and not the entire system (i.e., a district, county, or city system).

Goal Measurement. Even given a "total", organization-wide MBO program, the major obstacle to MBO in the educational context is goal measurement

(i.e., the "criterion problem"). The types of goals typically set by teachers and/or school systems lend themselves to be evaluated in one of two ways. Goals are either evaluated by ratings (principal and/or student) or pupil gain scores on standardized achievement tests. Both means of determining goal attainment have problems which shall be dealt with in turn.

A large number of studies have been conducted which used standardized achievement test scores as the criterion to measure teacher effectiveness (e.g., Anderson, 1954; Brophy, 1973; Ebel, 1973; Heil & Washburne, 1962; Wolfe, 1977). The educational literature has expressed mixed opinions as to the appropriateness of using such measures. Ebel (1973) defends the use of standardized tests as a basis for judging the effectiveness of a school system, stating that the information provided by these tests help school boards focus their attention on the weaknesses of the school system, thus allowing more immediate and purposeful remedial action. Most of the empirical data, however, would not support Ebel's position. Soar and Soar (1977), Brophy (1973), Hall (1964) and others argue against the use of achievement test scores.

Standardized Test Scores as Criteria. Airasian (1979) states that standardized achievement tests have been criticized on three basic grounds: 1) those characteristics intrinsic to the test, because of what they are, or what they measure; 2) those characteristics not intrinsic to them, what they aren't or don't measure; and 3) on the basis of what they do, their perceived effect on pupils, teachers, and the educative process in general. Airasian notes that those who argue that it is a misuse of tests to evaluate teachers' performance on the basis of their pupil's standardized achievement test

performance have done so by mentioning all of the characteristics cited above. For example, typical comments voiced by those in opposition to the use of standardized achievement tests are that such tests measure only a small sample of what teachers try to do (intrinsic factor), do not measure specific classroom goals (extrinsic factor), and to use such tests in this manner will mean that test content will dominate the classroom curriculum.

In reply to the criticism that standardized achievement tests do not match the specific objectives of an instructional program Cronbach (1963) offers the following. An ideal evaluation would include measures of all types of proficiency that might reasonably be desired in the area of question, not just the selected outcomes to which the curriculum directs attention. If one wishes to know how well a curriculum is serving the national interest, one should measure all outcomes that are worth striving for. Standardized achievement tests which sample a broader range are more likely to help answer the question of the adequacy of a particular curriculum than are tests tailor-made to the specific instructional objectives. While Cronbach's point regarding the usefulness of standardized tests for evaluating a curriculum are well taken, he does not speak to the issue of using such tests for the purpose of evaluating teachers. A serious misuse of these tests occurs when a single achievement test score is used as the sole criterion in an MBO or accountability program for making important personnel decisions.

Klien (1971) summarized the major problems of using standardized achievement tests as criteria as follows: 1) questionable test validity; 2) poor overlap between program and test objectives; 3) inappropriate test instruction and directions; and 4) confounding designs and formats. Averch et al.

(1972) add that such tests are too often misused or misinterpreted. Airasian (1979) reviews some of these misuses. In addition, when using standardized achievement tests, effectiveness or goal measurement is typically accomplished by the use of gain scores. Gain scores, however, produce a very biased estimate of true gain (see for example, Harris, 1963; Nunnally, 1978).

The empirical evidence appears to support those opposing the use of standardized achievement tests as criteria of teacher effectiveness.

Rosenshine (1970) attempted to ascertain the external validity of student achievement scores as a measure of teaching effectiveness. He noted that in only one of the nine studies reviewed was a correlation coefficient as high as $r = .50$ found while all the rest were at or below $r = .35$. He concluded that residual achievement gain scores cannot be used as measures of teaching success with any confidence.

Shavelson and Russo's (1975) review also questioned the ability of achievement tests to accurately assess teacher effectiveness. The authors note the consistent conclusions of previous research as to the "instability of teacher impact on student outcomes" (p. 29), which may be attributable to the instability of teacher behavior, an area that has not been assessed in its own right. They point to the problem of the lack of an operational definition of teacher effectiveness, and to strengthen their point, draw out four qualitatively different definitions from past research. Another problem they cite is the "collective display of stability coefficients which vary unsystematically in magnitude" (p. 177). They examined the studies in the context of the three dimensions of external validity as proposed by Snow (1974), and not one of those studies met the criteria on any one of those

three facets. The authors concluded that the instability of the data "makes generalizability in an educational context tenuous" (p. 38) and also recomended against the use of student achievement scores as measures of teacher effectiveness.

Brophy (1973) attempted to observe teacher efficiency scores (as plotted graphically over a three-year time span) as predictors of future success, as measured by pupil gain scores. He was able to observe 51% linear trends (linear consistency, improvement, or decline) from among the 145 teachers who taught at the same grade level for three or more years. Because of the moderately low consistency found, Brophy also warned against the "indiscriminant use of student gain on general achievement tests for assessing teacher accountability" (p. 252).

In a classic study, Barr et al. (1935) tested the validity of several measures of teaching effectiveness as measured by gains in pupil achievement. Reported correlations between the ninteen measures of teaching ability and pupil achievement were "uniformally low and insignificant" (p. 137). Independent variables included a composite of seven commonly used rating scales, a composite of six tests of teaching ability, and a composite of all 19 variables. The inconsistent data also failed to indicate a preference between the several measures of efficiency, or to validate any of those "commonly-used" teacher rating scales. The authors concluded that either the instruments chosen were externally invalid or the Stanford Achievement Tests were inadequate as criteria for gains in pupil achievement.

Anderson (1954) correlated pupil, supervisory, and self-ratings with mean pupil gain scores. Using a series of regression models using pre-test

scores, squared pre-test scores, pupil sex, and teacher effects variables, he also attempted to predict pupil gain scores. The rating scales had no significant effects as was the case with the regression equations. Anderson concluded that the results did not yield sufficiently high correlations to justify the use of residual gain scores as measures of teaching effectiveness.

Soar and Soar (1977), in their discussion of the problems inherent in using test gain scores as criteria for teacher effectiveness, state that unless individualized background factors (I.Q., socio-economic background) are considered, the criterion is useless and invalid. They state that as much as 80% of the variation in class averages on those tests is due to factors uncontrollable by the teacher. The authors state that differences exist in a teachers' success in instructing students on simple level items (requiring rote memorization only) and more complex items (requiring abstract thinking), and that this distinction has yet to be the focus of research.

Chase (1974) observed that standardized achievement tests measure a restricted sample of a child's educational behavior, and also tend to be overvalued in the assessment of academic performance. The tests, he claims, overstandardize the curriculum by identifying certain topics as more essential, which constricts individual initiative and creativity of pupils and teachers. He maintains these variables are the key to pupil educational growth and maturity. Ingenkamp (1977) notes that when noncognitive factors are used, they are used almost exclusively as additional predictors of cognitive performance. The consistent finding however, that these noncognitive factors do not relate to cognitive performance has resulted in the premature conclusion that their significance is limited. Ingenkamp states that noncognitive

factors such as reduction of school anxiety, the forming of relevant motivation and interests, and education toward emotional stability, all constitute important aims in education, whether they manifest themselves in the improvement of cognitive performance or not. Another problem according to Chase (1974) with the tests themselves is the subtests, which the author sees as having too few items to hold up to minimum validity standards as well as they should.

It appears that the most "countable" results in terms of educator effectiveness are replete with problems of measurement and interpretation. Empirical tests have for the most part not supported the use of achievement test scores as measure of teacher effectiveness. Their continued use has been attributed to the fact that they have face validity (Madaus, Airasian, and Kellaghan, 1980) and provide a "politically" attractive index on which to base accountability decisions (Airasian, 1979). Ebel (1973) however, points out that many of the empirical studies have been methodologically flawed. At this point we can only conclude that the use of standardized achievement scores in an MBO, as well as a rational-goal approach to school effectiveness, has great potential for problems of interpretation.

It has been suggested by Madaus et al. (1980) that to be considered valid, outcome measures in studies of school effectiveness must tap achievements, skills, processes, and learning that are specific to school instruction. If the outcome measures used to assess school effectiveness are primarily surrogates for home background and general ability, rather than measures of school-specific achievement, inferences about the differential effectiveness of schools are invalid, since the measures tell us little

about what is taught in schools. To deal with the relative insensitivity of standardized norm-referenced achievement tests for detecting school-specific pupil achievement, some have advocated the use of criterion-referenced tests (Cronbach, 1971; Glaser & Niko, 1971). It should be noted at this point that the same achievement test can be criterion-referenced test or a norm-referenced test (Mehrens & Ebel, 1979). The distinctions refer only to the interpretation of the test, not the test itself or the way it is scored. A norm-referenced interpretation adds meaning to a person's score by comparing it to those of other individuals in a specified group. A criterion-referenced interpretation, on the other hand, adds meaning to a person's score by comparing it to some specified criterion of proficiency. This latter interpretation appears more compatible with the goals of accountability programs and seems more appropriate for measurement for use in an MBO or rational-goal approach. It should be noted, however, that not everyone would agree with this recommendation. Shepard (1980) states that the standards used as criteria for criterion-referenced tests are arbitrary cutting points along a continuous performance scale which impose an artificial dichotomy. This dichotomy serves to obscure performance information about individuals along the entire continuum. In addition, since it is not possible to have an objective, nonjudgmental, standard setting process, Shepard recommends that standards not be used to interpret test data regarding the value of educational programs. However, by its very nature, decisions regarding effectiveness must always contain a judgmental component. The practitioner is also reminded of the other limitations in the use of test scores for effectiveness decisions, particularly the use of one index as a sole criterion and the fact that the tests are typically not

measuring school-specific factors. While a strong case can be made for the use of a criterion-referenced interpretation as opposed to a norm-referenced interpretation of standardized tests, the question still remains as to whether achievement test scores are an appropriate measure of school systems effectiveness. The data, at this point, seem to indicate that they are not.

There is one final criticism of tests pertinent to this discussion. While tests may distinguish between those pupils who have learned something and those who have not, they provide no information as to how or why the learning did or did not take place. To assess the how or why the most frequently used form of teacher evaluation is a rating done by either the principal or the students (e.g., Critenden & Norr, 1975; Frey, Leonard, & Beatty, 1975; McNeil & Popham, 1973; Scott, 1975; Watson, 1974). As with achievement test scores, however, there is little consensus in the educational literature regarding the appropriateness of their use. Scott (1975) and Watson (1974) view student ratings of teachers as constituting one of the most credible indicators of professional performance available and view the information obtained from evaluations as being valuable and necessary. Watson (1974) in one study found that at the high school level, teacher feedback from students was more effective than feedback from supervisors in leading to a positive change among teachers. This supports another criticism of MBO. MBO was not originally designed to be used as a performance appraisal strategy. It is best when used as an organizational development tool, the benefits of which are supported by Watson's (1974) findings.

Although used frequently, many have questioned the validity and fairness of supervisory ratings of teachers as presently conducted (Crittenden &

Norr, 1975; Marsh, 1979). Robinson (1978) notes that most evaluators (i.e., principals) have had no training in observational techniques. McNeil and Popham (1973) in a review of the educational literature, criticized supervisory ratings as being unsatisfactory on several grounds, including the confusion of a teacher's personality and staff relations with his/her teaching ability. In addition, teacher ratings, like any other employee ratings, are plagued by problems of halo, central tendency, leinency, etc. A thorough review of the errors associated with performance ratings is presented by Landy and Farr (1980). MBO typically calls for the measurement of countable, quantifiable results (e.g., achievement test scores) in order to avoid some of the problems in human judgment. With this approach, however, there is the distinct possibility that the more "process-oriented" variables will be ignored because they do not yield easily "countable" results. Exclusive attention to "end results" such as student achievement while ignoring teacher behavior could prove disastrous in terms of unmeasured, long range criteria such as student attitudes toward learning. While teacher ratings can serve as a vehicle to rate teacher behaviors, they are typically not used as such. This is disheartening in light of Rosenshine and Furst's (1971) findings based on reviewing 50 studies of classroom instruction. These authors found that growth in cognitive achievement appears to be promoted by 11 classes of teacher behavior (e.g., clarity of presentation, enthusiasm, avoidance of strong negative criticisms). In addition to behavioral components, Hanushek (1977) stresses the importance of teacher's personality characteristics as they relate to student performance.

As a practical matter, there are a number of problems inherent in goal

Tebay (1973) stresses the importance of employee participation in goal setting. The question thus arises as to whether individuals are truly interested and/or willing to participate in the goal setting process, and if they are, are they capable of such participation? According to Gray and Burns (1979), a major reason why individual teachers are reluctant to participate in goal-setting, and why MBO doesn't work in schools, is because performance appraisals in school systems are not tied to the reward system in any way. Levinson (1970) rejects the model outright, by stating that since it is not tied to the reward system it will subvert its own intentions. Salary increases, advancement, and lay-offs occur as a function of collective bargaining and tenure. There are few, if any, penalties for mediocre performance.

It would certainly be beneficial if performance was tied to the reward system, particularly at the elementary school level where MBO appears to be most appropriate. Much like the rational-goal model discussed earlier, MBO provides a good framework within which to assess the effectiveness of teaching basic skills at the elementary school level. Goals can be easily established (e.g., increase mathematical ability from a second to a third grade level) that are compatible with all parties concerned, and their attainment can be measured reasonably well by criterion-referenced interpretations of standardized achievement tests. The next question becomes, is goal measurement alone indicative of effectiveness? The missing component that would make MBO and the rational-goal approach complete as models of school system effectiveness, would be a concern with behaviors as well as outcomes. The functional model presented next incorporates this missing

component.

The Functional Model

This model, as delineated by Cunningham (1977), states that effectiveness is determined by the social consequences of an organization's activities. The crucial question is: how well do the activities of the organization benefit its client groups? In terms of school systems these groups may be parents, students, taxpayers, other schools, etc. The organization sets certain goals and seeks to attain them through satisfaction of the client groups. These goals should be precisely defined, and the activities necessary for their accomplishment should be delineated. Appraisal of organizational effectiveness considers whether these activities are functional or dysfunctional in achieving its goals. Unlike the other approaches there is an emphasis on "means" in addition to "ends".

The implication, however, is that this functional-dysfunctional dimension applies only to the organization (i.e., the activities are only assessed in relation to the organization, not the client groups). Thus, the organization's activities would be assessed as good or bad only in regard to their effect on the organization itself. Generally, effects of these activities on relevant client groups are not considered. This seems to be a fatal flaw in a theory whose frame of reference is the external society.

Michael and Metfessel (1967) present a model of educational evaluation that is very compatible with the functional perspective. Their model attempts to account for the effects of organizational activities on relevant groups external to the immediate organization. Their model calls for: (1) the evaluation of programs involving members of the total school community such

as lay individuals and groups, professional personnel, students and student groups; (2) the construction of a hierarchical list, beginning with general goals and going to specific objectives which are operationally defined that allow relatively objective measurements whenever possible; (3) a translation of the specific behavioral objectives into a form that is communicable and applicable; (4) the development of criterion measures which reflect program effectiveness in terms of the objectives; (5) the periodic use of these measures to assess behavioral changes consistent with the objectives; (6) the analysis of data generated by these measures with appropriate statistical methods; (7) the interpretation of these data in terms of certain judgmental standards to derive information about the direction of growth, students' progress, and the effectiveness of the total program; and (8) recommendations which are based upon repeated cycles of the evaluation process and which provide further implementation and modification of the broad goals and specific objectives.

The authors also present multiple criterion measures for the evaluation of school programs, such as: (1) indicators of status or change in cognitive and affective behaviors of students using standardized instruments such as achievement tests, self-inventories, etc.; (2) indicators of status or change in cognitive and affective behaviors of students using informal or semiformal teacher-made instruments such as incomplete sentence techniques, interviews, peer nominations, etc.; (3) indicators of status or change in student behavior using means other than the above such as absenteeism, attendance, anecdotal records, etc.; (4) indicators of status of change in cognitive and affective behaviors of teachers and other school personnel such as attendance, published

articles, elective offices, etc.; and (5) indicators of community behaviors such as alumni participation, attendance at special school events, parent-teacher conferences, etc. With this paradigm, more consideration is given to "means" (i.e., behavioral measures) than with other goal-oriented approaches (e.g., Cunningham, 1977). However, as with other goal-oriented approaches the problems of goal measurement still exist and are not discussed by Michael and Metfessel (1967).

Goal Attainment Scaling. Goal attainment scaling (GAS) is another popular goal-oriented approach that fits within the functional model of organizational effectiveness. Introduced by Kiresuk and Sherman (1968), GAS is a technique designed principally to assess individualized goals in specialized programs. There have been several applications of GAS to the educational domain (Cytrynbaum, Ginath, Birdwell, & Brandt, 1979) with concentration on programs in special education. Kiresuk and Sherman (1968) list the major prescriptions for GAS as: (1) change goals must be set independently of the process (goal setter and deliverer should be independent); (2) clients should be randomly assigned to treatment conditions; and (3) the assessment of goal attainment should be done by independent raters. Cytrynbaum et al. (1979) note that most of the studies evaluating applications of GAS for school programs have deviated from these essential requirements (see also Calsyn, Tornatzky, & Dittmar, 1977). Cytrynbaum et al. (1979) conclude that the popularity of GAS is unjustified given their thorough scrutiny of results. The authors recommend further research directed at a clarification of critical and influential parameters affecting the GAS approach. It should also be pointed out that Cytrynbaum et al. cite

methodological and design factors that preclude definitive conclusions regarding GAS.

Fremont (1975) presents an excellent discussion of the appropriateness of the functional model for public organizations. According to his analysis, the major features of the functional model are: (1) goal attainment; (2) adaptation; (3) integration; and (4) pattern maintenance. Regarding the first feature, Cunningham (1977) and others discuss goal attainment in terms of a set of specific output values that are important for key client groups. The "adaptation" function deals with a determination of resources that are necessary to achieve the stated goals. Problems such as conflicting goals of client groups (i.e., taxpayers, parents, students) and measurement issues regarding goal attainment (Averch et al., 1972; Hanushek, 1977), however, may attenuate the appropriateness of the model for school effectiveness. The "adaptive" aspects of goal attainment such as resource procurement, management, and budgeting also represent conflicts between client groups and organizational members (e.g., teachers). The "pattern maintenance" function involves a reduction in conflicts across client groups and within the organization (e.g., increasing teacher job satisfaction). This function, however, is only considered as a "means" toward goal attainment. Thus, given conflicting "client" goals, even "pattern maintenance" functions will engender conflicts across client groups and within the organization. The lack of empirical support to justify "pattern maintenance" functions for goal attainment (e.g., better facilities will result in higher pupil achievement; higher teacher pay will result in higher pupil achievement and higher non-cognitive skills) predicts constant strains for the integrative aspects

of the organization's functions and in its relationships with client groups.

One of the major shortcomings of the functional model as it has typically been applied to school systems is its preoccupation with "end-states" and its consideration of "adaptative" and "pattern maintenance" functions only as "means" toward the "end-states." The functional model is not concerned with "means" in and of themselves, rather it emphasizes "ends". The functional model than, in practice, is equivalent to an MBO or rational goal approach. Thus problems of goal articulation and goal measurement that beset other goal-oriented approaches also affect the efficacy of the functional model for educational systems.

A major criticism of the functional model is discussed by Perrow (1972) who questioned the unidirectionality of the approach as conceptualized by Parsons (1960) and others. It is the nature of static research, he states, that implies the organization's activities affect its client groups and not vice versa. In the educational context, not only does the unidirectionality assumption make little theoretical sense but there is fairly strong empirical support to refute it as well (e.g., Averch et al., 1972; Hanushek, 1977).

At a conceptual level the functional model appears to contain all of the necessary components of a complete effectiveness model. Unlike the rational-goal model and MBO, it incorporates both "ends" and "means". The functional model is concerned with the effect of a school system's activities on its client groups. Unfortunately, the model as discussed by Cunningham (1977), has not been applied to school systems. Where the model has been attempted, it typically concentrates on "ends" at the expense of investigating "means". While appropriate at a theoretical level, practitioners should be

aware that in applying the functional model to school systems, they must contend with the problems of goal-approaches (e.g., goal measurement) plus the added difficulty of measuring behavior (e.g., which behavior should be measured? how? and by whom?). Given the difficulties involved in the latter it is not surprising that most applications of the functional model have focused exclusively on outcomes in lieu of teacher behaviors.

Thus, for the practitioner concerned with assessing the effectiveness of a school system, the functional model offers the optimal solution focusing on outcomes and behaviors. The application of this approach to an ongoing school system, however, would be difficult and time consuming. It would be time consuming in the sense that someone, be it fellow teachers or principals, would need to take the time to observe teachers so they could evaluate their behavior. It would be difficult, in the sense that (1) a teacher's activities are not easily measured or evaluated and (2) different teaching styles may yield similar results. Who is to say what method of teaching is correct? The greatest difficult, however, is that ideally practitioners would like to relate these observed behaviors to objective outcome measures, which brings us full circle, back to the problem of validity for standardized tests which has already been discussed.

The Cost-Benefit Model

Cost-benefit models have generally been directed at the evaluation of the relative effectiveness of different training programs. Such an evaluation is most often conducted in a goal-centered framework (e.g., Rivlin, 1971) and thus addresses the question as to whether an organization is accomplishing its goals with procedures that are cost-effective relative

to other procedures (cf., Wolfe, 1977). The model implies that there are alternative courses of action available to the organization, and that they can be compared. This is not usually the case. Efficiency is the central variable in the cost-benefit model as with the systems-resource model (to be discussed in the next section). Therefore, the general problem with measures of this construct addressed in the context of the systems model applies to cost-benefit models as well (e.g., Fincher, 1972).

Simply stated, the goal of the cost-benefit model is to optimize the ratio of costs to benefits. There are two assumptions implicit in cost-benefit models: (1) that the components of both the numerator and denominator can be empirically, intuitively, or rationally reduced to single composite scores; and (2) that the resultant composite scores have interval scale properties (Campbell et al., 1974). Both of these assumptions have been questioned, particularly in applications to school systems (Hatry, 1970; Rivlin, 1971). Additionally, this model necessitates the use of outcome variables that are more distant and less under control of the organization being studied (Campbell et al., 1974).

The major advantage of the cost-benefit model as applied to school systems is that it provides an analytical framework for evaluating various educational strategies. Research has been directed at refining measurement strategies for both inputs and outputs (cf., Chase, 1968; Glennan, 1972; Mangum, 1967).

Fansel and Bush (1970) provide an excellent example of a cost-benefit model applied to a social system; specifically, a health care system. They also present a method for converting ordinal data to interval data. In their

study, effectiveness was defined as the degree of function or dysfunction in the affected population. This continuum was defined in terms of a person's ability to carry on his/her daily activities. Using the paired comparisons scaling technique, subject matter experts derived interval scales from judgments of the ordinal descriptions. The authors presented some data on a tuberculosis control program to illustrate cost-benefit analysis.

As applied to the educational context, Hatry (1970) maintains that cost-benefit analysts do not consider goals or objectives with an open systems perspective. Thus, constituencies who may be directly affected by decisions from cost-benefit analysis may have little or no input into the setting of operative goals on which the cost analysis is based. Furthermore, he also questions the overemphasis on measures of efficiency to the detriment of measures of effectiveness. The majority of cost-benefit studies use pure cost measures such as expenditures per pupil (e.g., Law, 1977). Other studies have used workload measures such as number of teachers per pupil or physical standards (students per square foot or per classroom). Airasian et al. (1979) refer to these as static variables, and note that while they are easier to manipulate, such proxy measures are only indirectly (if at all) related to measures of effectiveness. A smaller group of studies have attempted to develop composites to several measures using a common metric such as dollars. Hatry (1970) classifies such measures as "hocus-pocus" in that they conceal the value judgments made by the analysts. Rivlin (1971), Hatry (1970), and Campbell et al. (1974) recommend a multiple criterion approach to effectiveness whereby all relevant constituencies (see Connally, Conlon, & Deutsche, 1980) assist in the specification of unambiguous objectives

and methods of assessing their attainment. Such criterion should not be made equivalent using "bogus" dollar figures simply for the sake of forming a composite criterion.

Wolfe (1977) provides an example of a cost-benefit approach that incorporates an operationalization of effectiveness that is more closely linked to the general objectives of educational systems. He states that his production function approach can help school policy-makers make rational decisions involving: (1) defining the goal(s) of the system; (2) collecting data relevant to each pupil; and (3) analyzing the information in a systematic way that relates inputs to defined goals of output.

Wolfe studied a large urban school system over three years to determine input-output relations for grades three-six. The major questions of his study were: what school and teacher characteristics will maximize pupil achievement growth and, combining school and teacher characteristics, what are the most effective inputs per dollar. The independent variables used in the study were: Teacher experience; National Teacher Exam score; school size (enrollment), class size; library books per pupil; rating of teacher's undergraduate instructor; educational level of teacher; and school characteristics (i.e., play-ground footage, date built, condition of school). The dependent variable was student gain scores on the Iowa Test of Basic Skills (grades three-six). The production function used in his analysis was:

$\Delta A = f(T_1T_1 \cdot 6SES, T_2, T_2 \cdot 6SES, \ldots, S_1, S_1, \cdot 6SES, \ldots S_n, S_n \cdot 6SES/6SES, P6)$

where ΔA is change in achievement score; $T_1 - T_4$ are teacher quality measures; $S_1 - S_n$ are school quality measures; 6SES is genetic endowment socioeconomic status; and P6 are peer characteristics.

The results supported the contention that a systematic evaluation of inputs in relation to outputs, combined with cost figures, can increase the efficiency with which educational resources are used. For example, Wolfe (1977) found that the cost per pupil was not directly tied to pupil achievement growth (as previously found; Hanushek, 1977). He concluded that current resources can be used more effectively by reallocation based on production functions as in his study. Wolfe (1977) also found that different expenditure patterns yield very different results and that the appropriate expenditure pattern could be selected as a function of the priority of goals from the multiple constituencies or the "dominant coalitions".

While Wolfe's approach to cost-benefit analysis is an improvement over many previous attempts to link educational costs to outputs, the major drawback of the study is that only one criterion (pupil gain scores) could be used in his operationlization of effectiveness. As previously noted, Ebel (1973), Nunnally (1978), and Soar and Soar (1977), discuss some of the problems with using pupil gain scores on achievement tests. In addition numerous writers have cited the importance of time frame on which the criteria are based (e.g., Mahoney & Weitzel, 1969; Yuchtman & Seashore, 1967). Effectiveness is often divided into categories such as immediate (or short-run), intermediate (middle-range), and ultimate (long-run). Static designs, which are characteristic of cost-benefit analysis, would assess only immediate effects and ignore the impact of such effects as intermediate or long term criteria. For example, the use of cognitive measures as output at T_1 could have deleterious effects on the same measures at T_2 or T_3. Also, the exclusive concentration on factors dealing with cognitive development could

result in a negative impact on non-cognitive development at T_1, T_2, or T_3. This could prove to be a serious problem at elementary educational levels given the recognized importance of non-cognitive development at this level and the implied linkage of cognitive achievement with resource allocation. The problem is similar to that which is characteristic of other goal-oriented approaches; that is, attention is directed toward measurable objectives at the expense of so-called non-measurable, non-cognitive components (e.g., creativity, abstract reasoning).

At this point a distinction between cost-effectiveness models and cost-benefit models needs to be made. While effectiveness requires the quantification of input costs such as teacher characteristics and facilities, the outputs or benefits do not require a cost estimate attached to them. Rossi, Freeman, and Wright (1979) use the example of a free textbook distribution program for primary school children that may have findings of cost-effectiveness expressed as "'x' total project dollars increased reading scores by 'y' amount". Thus, the textbook distribution program can be directly compared to some other form of intervention, the goals of which are also aimed at increasing reading scores. Comparisons and rankings of choice can therefore be made on the basis of magnitude of benefits versus costs.

Related to the problems of the criterion (benefit) selected in a cost-benefit analysis is the underlying accounting perspective to be adopted. For example, for educational programs, there are three accounting perspectives that could be used; the individual, the society-at-large or the administrating government. Garms (1971), for example, in assessing the costs of a remedial education program, identified individual costs such as tuition

rates for college attendance, earnings foregone while attending college, and additional expenses while in college. The social costs identified were opportunity costs (gross of taxes) and project costs. The government costs were taxes lost and any project costs incurred. Rossi et al. (1979) and Glennan (1972) provide several examples illustrating that the same components comprising a cost-benefit model may enter into the equation as a cost from one perspective and a benefit from another. It follows then that any ratio derived from the model will be a direct function of the accounting perspective that has been adopted. Paul (1973) in an analysis of management education, presents an example of the extreme differences in ratios that result from a change in the accounting perspective. Rossi et al. discuss the details involved in such an analysis.

While the different accounting perspectives may result in different cost-benefit ratios, the adoption of an effectiveness model where benefit factors other than, for example, earnings improvement can be considered might increase the compatibility of different perspectives. Furthermore, the derivation of cost-effectiveness models from the different perspectives could serve as the basis for discussion by the multiple constituencies or "dominant coalitions" affecting the system. Such an approach would be more desirable than that which is typically done when relative comparisons must be made between various educational programs in order to decide budgetary matters (Pellegrin, 1976). Of course, while this sounds like a better approach to decision-making, it necessitates the pricing of variables that are not easily priced. What, for example, is the price of the improved cognitive achievement of the learning disabled versus that of the high

achievers, or the development of physical abilities versus art or music appreciation? Asking people what they are willing to pay for improved cognitive development of their children is one proposed method of pricing (Posavac & Carey, 1980) but is inappropriate whenever the majority of constituencies are not directly affected by certain decisions (for example, pricing for special education programs that do not directly affect one's own children).

Another problem with cost-benefit approaches for school systems is that the preoccupation with economic analyses (e.g., earning power) tends to devalue the importance and quality of human life. Little consideration, for example, is given to factors such as life satisfaction other than how it relates to earning power (Abt, 1977). This problem is related to the difficulty of pricing certain benefits. Oftentimes when prices are assigned to benefits, the decisions for pricing are arbitrary or purely speculative (e.g., Levin, 1975).

Macy and Mirvis (1976) describe the development and implementation of a standardized set of definitions, measures, and costing methods for various behavioral outcomes related to the quality of work life. In a longitudinal study of manufacturing and assembly plants, the authors derived dollar figures for four variables related to member participation (absenteeism, tardiness, turnover, and work stopages) and six variables reflecting role performance (productivity, product or service quality, grievances, accidents, job related illnesses, and unscheduled downtime). No studies could be found in the educational literature that assigned dollar estimates to as many salient variables. Additionally, no attempts could be found where administrator, teacher, or

student attitudes are assigned dollar values (see, for example, Mirvis & Lawler, 1976). Thus, in terms of a large scale model of school system effectiveness, applications of a cost-benefit approach have been largely incomplete. However, when educational decisions are closely linked to dollar costs as benefits, cost-benefit or effectiveness approaches should prove to be a valuable input. It must be stressed, however, that for the practitioner concerned with increasing the effectiveness of a school system, cost-benefit analyses are only one factor that should enter into an effectiveness model. There are additional factors (i.e., inputs) that are difficult if not impossible to operationalize in dollar amounts (e.g., teacher behaviors, student attitudes, organizational climate). Thus, a cost-benefit analysis by itself is not appropriate to assess school systems effectiveness but could be beneficial when used in conjunction with other approaches.

Human Resource Accounting. Other than as discussed above in terms of earned income and relative program cost/benefit ratios, there have been few attempts to apply principles of human resource accounting (HRA) to the measurement of educational efficiency. The theory underlying human resource accounting is that the value of humans to an organization can be determined. The different costing methods are surrogates of the actual value of the human organization.

Flamholtz (1974) has made extensive studies of the value of an individual to an organization, or the individual's expected realizable value. This theory recognizes several elements which give people value as human resources because they represent potential services to the organization (e.g., skills and activities of persons upon entering the organization, reward structure

of the particular role, promotability, and turnover). This theory identifies major factors which constitute a person's value and organizes them into a logical framework.

With regard to assessing the value of the total human organization, Likert (1973) has developed a basic theoretical approach. Likert's work actually could be considered the basis for the current interest in HRA. His human resource accounting model included causal, intervening, and end-result variables. He stressed in his model the interactive nature of all variables within his model with managerial leadership conceived as the causal variable. He later included organizational structure as a causal variable as well as discussed the influence of this variable on intervening variables such as organization climate, peer leadership, group processes (planning, coordinating, decision-making, etc.), and subordinate satisfaction. The intervening variables in turn affect the end-result (i.e., the total productive efficiency). Likert posits differences in the magnitude of the relationships between the structural variables and the intervening variables. Such an approach could also be presented in the context of the managerial process model of effectiveness.

Likert (1973) suggests that the causal and intervening variables can be measured by existing socio-psychological tools. The causal human organization scores then can be related to unit "production" costs through placing dollar values on changes. In the educational domain, however, attempts to relate teacher training costs, for example, to output measures such as student achievement have been largely fruitless (e.g., Averch et al., 1972; Hanushek, 1977). The educational studies, however, have been far less

sophisticated in their measurement strategies and make no attempt to relate human resource accounting principles to teacher or administrator management. For example, using past human resource costs, standard costs for the acquisition of a new teacher or principal in a certain position could be developed. These could be used to evaluate the efficiency of teacher recruiting and hiring much as standard costs are used as guidelines in a production department. The areas where costs are greatest could be determined and possibly reduced. If costs need to be reduced, it will be an informed reduction rather than an arbitrary cut back as is usually the case (Pellegrin, 1976). Standard costs are always useful for the practitioner in budgeting. If the number of positions needed to be filled is available, an appropriate budget for personnel acquisition can be determined.

Capitalizing training costs would decrease the tendency of management to view training as a luxury item. Training costs could be accumulated in the asset accounts of the individuals who benefited from the training sessions and matched later with output it helped to produce. School administrators would be encouraged to upgrade their continuing education programs since the programs will be evaluated by long-range results (increased performance) rather than short-range increased costs. The evaluation would include a return on investment figure being computed for human assets as well as physical and financial assets. Additionally, HRA data is useful in evaluating turnover. If excessive turnover occurs, it will be reported as losses on the budget statement. Basset (1972) presents an analysis of employee turnover that has application to problems such as "high quality" teacher turnover and "low quality" tenure.

Tomassini (1977), in a study of layoff decision preferences comparing the decisions made with and without access to HRA information, tentatively concluded that HRA cost estimates caused different managerial preferences in the decision to layoff personnel. In this case the HRA costs included costs to rehire, and to hire and train new workers to replace those who did not return. Not all of the subjects with access to HRA data made the same decision but none appeared to have ignored the HRA data.

Several models have now been presented for HRA (Craft & Birnberg, 1976). Further research should concentrate on empirically testing these models. Before school systems implement HRA methods, they need to see whether the benefits will exceed the costs. One advantage for the practitioner is that implementing the system need not be organization-wide at first. AT&T, for example, has used HRA in measuring the cost of employing and developing toll directory and assistance operators (Stone, 1972). This was a problem area for the organization yet one that could be objectively measured in terms of costs. A similar approach could be adopted for educational systems with a concentration on one major personnel segment, such as teachers. If the system is effective it can then be expanded to include other segments of the school system.

While there are numerous journal articles that discuss HRA and provide the proposed advantages and uses, there have been no pure applications of HRA to educational systems. The next logical step should be implementation on a limited basis and a reporting of the results. If HRA is to be used in a cost-benefit or effectiveness framework for school systems, it is necessary for educational management to be familiar with HRA and its uses

and to test and implement its principles where appropriate. It is also necessary that practitioners be aware of the fact that while HRA can provide a valuable input for school systems, one should be cautious not to base effectiveness decisions on HRA alone. It is recommended that HRA be used in conjunction with the functional-model of effectiveness. Combining these two models would allow the practitioner to evaluate "means," "ends," and costs.

Systems Models

The second major view of organizational effectiveness is the systems view (Campbell et al., 1974). This model regards the organization as a natural whole or system. Realization of goals is but one component of the measurement of effectiveness in this model. Rather, the organization adopts the overall goal of maintaining its viability or equilibrium without depleting its internal or external resources. The organization is seen to be in a constant state of flux, with its structures spontaneously and homeostatically maintained (Ghorpade, 1971). Etzioni (cited in Steers, 1976) argued that this approach was superior to the goal-oriented approaches because it focused on the system at work. Georgopoulos and Tannenbaum (1957) also maintain that definitions of organizational success must consider not only the objectives of an organization but also the mechanisms by which it maintains itself and pursues its objectives. The focus is thus on "means" as well as "ends."

The systems approach is based on an underlying organismic model that stresses the interdependence of the various subsystems. To assess effectiveness under a systems viewpoint, one should ascertain the extent to which

an organization is internally consistent, whether resources are being judiciously acquired and distributed over a wide variety of coping mechanisms, and if the organization is adjusting to a constantly changing environment (Campbell et al., 1974). This is a very dynamic perspective. Indeed, Ghorpade (1971) states that systemic criteria of organizational effectiveness are derived from conceptualizations of "needs" experienced by the organization as a "living" social system. Needs refer to requirements that organizations have to meet in order to survive. Conceptualizations of needs and their criteria take many forms (Steers, 1976). Bennis (1966) in fact conceptualizes the problem from the perspective of mental health.

The reliance on goals, however, is still very much in evidence in some systems models. Ghorpade (1971) notes that Parsons and Etzioni, two staunch advocates of the systems approach, both insist on linking effectiveness with goal attainment. They view the organization as one subsystem in a larger social structure. Organizational goals emerge as outputs for the society. Judgments of organizational worth are made in terms of the quality and relevance of the organization's outputs toward assuring the growth and survival of some other system, usually society at large. However, social accountability, an important issue for school systems, has not been formally recognized in any theoretical system or framework. In addition, any systems model that incorporates the concept of goal attainment is subject to the same problems noted earlier about goal-centered approaches.

Presented in Table 2 are five of the more prominent systems approaches in the literature on organizational effectiveness. The table summarizes these models with respect to their applicability to school effectiveness.

A detailed discussion of these models in terms of their theoretical underpinnings, empirical support (where available) and implications for the assessment of school effectiveness will follow.

Insert Table 2 about here

The Systems Resource Model

Systems-resource models define organizations as several subsystems interacting with each other (Etzioni, 1960; Katz & Kahn, 1978). Effectiveness is defined as the degree to which the subsystems work harmoniously to satisfy needs of constituencies. Cunningham (1977) states that the key to effectiveness is a balanced distribution of resources among the various subsystems, not maximal satisfaction for all subsystem needs. Katz and Kahn (1978) write that the "ultimate justification of an organization's existence and its claim on scarce resources is its throughput, its transformation of energy and material into forms and locations valued for reasons and by individuals outside the organization. The more the throughput, the better its quality, and the less consuming the transformation process, the more we call the organization effective" (1978, p. 246). Yuchtman and Seashore (1967), who also adopt an open systems perspective of organizations, define effectiveness in terms of an organization's bargaining position or the ability of the organization to exploit its environment in the acquisition of scarce and valued resources. Thus, the only relative comparisons that can be made between organizations is in the sense of ability to compete for valued resources. It should be noted that Yuchtman and Seashore (1967) emphasize "ability" to compete rather than the measurement of actual acquisition of resources. It is of course difficult to conceive of how such "ability" could be assessed except through the acquisition of resources.

Price (1972) and others have argued that the acquisition of resources is the operative goal of the organization. Katz and Kahn (1978), however, maintain that the "throughput" represents the primary goal of the organization that has been set by the relevant constituencies of the organization. Effectiveness can then be defined in terms of the attainment of these goals given certain specified constraints. Therefore, for the practitioner, it would be possible to specify a single criterion of effectiveness (e.g., achievement test scores) given these specified constraints. With respect to educational systems, constraints such as community characteristics and political attitudes must be taken into consideration when measuring effectiveness. Barro (1970) also maintains that adjustments must be made in teacher characteristics across districts because such characteristics cannot be controlled by administrators. However, it is difficult to discern how factors under the control of administrators such as teacher recruitment, personnel selection, resource allocation and working conditions would not have an effect on teacher characteristics. Other obvious constraints that may influence measures of effectiveness for school systems are the characteristics of the population of students (entry-level), the market for all school personnel, the physical facilities available, and the resources available. Yuchtman and Seashore (1967) would probably maintain that for school systems variables such as tax base, or percentage of budget represent the bargaining component of effectiveness and, as such, are not constraints on effectiveness but rather measures of it. Thus, in their conceptualization of effectiveness, a bargaining variable such as percentage of total agency budget might be appropriate for comparisons between two school districts.

While there is certainly a goal-orientation to "throughput" with the systems-resource model, there is a concomitant concern for the maintenance of internal day-to-day activities and the ability of the organization to integrate relationships between its various subsystems (Cunningham, 1977; Etzioni, 1975). Etzioni (1975) writes that the "starting point for this approach is not the goal itself but a working model of a social unit capable of achieving a goal . . . it is assumed a priori that some means have to be devoted to such non-goal functions as service and custodial activities, including means employed for the maintenance of the unit itself" (Etzioni, 1975, p. 135). Thus, it is the systems-resource perspective that considers means as well as ends in its conceptualization of effectiveness. A pattern of interrelations among the various organizational subsystems is defined as the effectiveness model. In this sense, then, the systems-resource model resembles the organizational development model, to be discussed later, in that a set of assumptions regarding organizational characteristics are defined a priori as assumptions predicting or correlated with effectiveness. To Katz and Kahn (1978) and other open system theorists, effectiveness is most dependent on an organization's efficiency. Efficiency is critical because the survival of an organization is a function of its negentropy or the ratio of imported energy to returned output. Thus, an output/input ratio of 1.00 would represent an ideal state of efficiency for a school system. For the practitioner, then, the logical questions to follow from this conceptualization are: (1) how much output do we get for a given input; (2) how much input must we invest to assure a given output; and (3) is it possible to compare the relative efficiency of two or more similar

school systems?

The output/input ratio will tell us how well a school system is using the energy (inputs) at its disposal and how much energic investment is required for a given unit of output. In an educational setting, energic investment would entail such variables as personnel costs, supplies, power, facilities, etc. (O'Donoghue, 1971). It is in the measurement of the output variables that problems develop. Very often, simply the number of students who have passed a certain academic hurdle or graduated is taken as the measure of output. With this conceptualization, comparisons could be made between two educational systems that are essentially comparable in terms of input variables with the exception that one system maintains a teacher-student class size ratio of 20:1 while the other is at 30:1. Using an output variable such as the number of percentage of "successful" graduates, it is clear that the system with a 30:1 ratio is relatively more efficient than the other system. While both systems are operating as designed, one system is relatively more efficient than the other.

One flaw in this notion of efficiency when applied to educational systems is that the relative superiority of the 30:1 system may be evident only so long as the output criterion is the sheer number of graduates. Perhaps the system with higher teacher-student ratios also is characterized by higher teacher turnover, absenteeism, or lower student achievement and satisfaction. Thus, we once again return to the old "criterion problem" with regard to the measurement of efficiency. If constituencies can agree on appropriate output measures and energic inputs can be reliably and adequately measured, then efficiency would be an appropriate operationalization of school

effectiveness. There is general agreement, however, that the measurement of both input and output variables is inadequate and the very inadequacies of this measurement will tend to reduce the efficiency of school systems. As decisions are made in areas such as resource-allocation on the basis of data, to the extent that the data are invalid or limited in scope, so also will be the decisions. As an example, while few educators would argue with the contention that the lower the teacher-student ratio, the greater the student achievement, empirical studies in this area for the most part are not supportive. Hanushek (1977) writes that "almost uniformly, educational production models show no consistent or significant relationship between achievement and expenditures per pupil (either instructional expenditure or total expenditure)" (p. 47).

Current attempts to refine the measurement of organizational input and output variables have focused on some form of dollar criterion to be assigned to variables previously unaccounted for. As mentioned earlier, Likert (1967) and Flamholtz (1974) have applied human resource accounting principles in order to assign dollar estimates to input variables such as employer experience, selection, training, etc. However, there have been few attempts to apply cost accounting principles to output variables in an open systems context (Popham, 1975). In Parson's (1956) analysis of social structure, a school system's functions and effectiveness can only be considered in terms of its contribution to the functioning of the next higher order of social structure. Thus, cost accounting principles could conceivably be applied to school output measures such as "drop out" rate, percentage of students who go on to college or technical school, percentage of unemployed graduates,

etc. Such output measures would get closer to answering the question of "how well the organization is doing for the suprasystem" (Katz & Kahn, 1978, p. 246). With this conceptualization of effectiveness, there are many similarities between Parson's (1956) functional model and Katz and Kahn's systems-resource model. However, Seashore and Yuchtman (1967) take the position that such an approach neglects the independence of the organization and its exercise of choice. With their perspective, maximization of return to the organization underlies the measurement of efficiency. However, to use a measure such as percentage of total agency budget in order to make relative comparisons between school districts ignores the ramifications of budget increases on budgets of other suprasystem agencies. Also, it would be virtually impossible to control for all extraneous factors that could affect such a measure (e.g., general state of the economy, tax base). Furthermore, in viewing a school system in the context of its suprasystem (e.g., the community, local government, etc.), it is probable that one particular level in the school system hierarchy may have incentives that prove to be disincentives at another level within the hierarchy. For example, the "cost savings" realized by one level (e.g., administration) may be represented as "cost avoidances" to a level above the school board (see also the discussion on cost-benefit models).

In summary, the major problem with the systems resource model as applied to school systems effectiveness is the development of an acceptable standard for efficiency that takes into consideration the interaction of school system effects with effects of other institutional subsystems within the organization's suprasystem (O'Donoghue, 1971). The problem is compounded

by the fact that issues related to efficiency such as resource allocation are closely linked to community political processes. Research indicates (e.g., Pellegrin, 1965, 1968) that in matters of resource allocation, issues specifically related to education receive lower priority than matters related to economic and governmental affairs. Pellegrin (1976) discusses how the "community elite" spend little of their time and effort ruminating over educational issues and yet they make decisions that often directly affect educational outcomes. In education, the most common measures of efficiency are student/faculty ratios, costs per faculty member, costs per student, and costs per square footage of space (Bowen & Douglas, 1971; Cameron, 1978; Hartmark, 1975, Meeth, 1972; O'Neill, 1971). While these operationalizations are certainly important in terms of indicating efficient uses of resources, they bear little relationship to operationalizations of the effective use of resources (i.e., producing outputs). Fincher (1972) concludes that it is impossible to assess both efficiency and effectiveness in education using the same criteria and that the emphasis should be on measures of effectiveness. The preoccupation of systems - resource models with the measurement of efficiency make the model largely inappropriate for educational systems.

The Functional-Structure Model

The functional structures model is probably best represented by Pennings' (1975) "structural contingency" model which states that organizational structure is in part a function of environment and technology. The environment is the many parties with which the organization interacts (e.g., parents, students, teachers, and tax payers). Technology refers to the organization's

mode of operation (i.e., the way it converts its inputs into outputs). This model proposes that organizational effectiveness is a function of the interaction of relevant technology and the organizational structure. Unfortunately, the model lacks both a methodology for assessing most environmental factors and hypotheses that would clearly pertain to educational systems.

While many of the internal structures that have been studied in the context of the functional-structures model (e.g., decentralization, specialization, formalization) have also been considered for educational systems, none have been closely linked to measures of school system effectiveness (Averch et al., 1972; Hanushek, 1977; Madaus et al., 1980; Mohr, 1973). Commenting on the structure research conducted in educational settings, Mohr (1973) has stated that few hypotheses can be found with regard to how certain structural inputs will affect school outputs. He states further that factor analytic research involving a variety of input variables has been largely fruitless.

A major limitation of the functional-structural model as applied to school systems is that it has been principally concerned with the satisfaction of basic needs that should probably not be the focus of a model of school system effectiveness (e.g., the security of the school system in relation to its environment; the continuity of policy-making). Furthermore, the evaluation criteria for this model have most often involved profitability, the acquisition of resources such as money, people, or goods, and investment in the system through goods or people (Cunningham, 1977; Selznick, 1953) criteria which are not particularly applicable to public school systems.

Grayson (1972) discusses the need to implement an educational technology which involves curriculum design and redesign based on a priori criteria. While his discussion resembles Penning's notion of the structural-technological interaction, the approach more closely resembles the systems-resource perspective proposed by Yuchtman and Seashore (1967), Katz and Kahn (1966), and others.

Miskel, Feverly, and Stewart (1979) maintain that there have been three common deficiencies in research on educational structure variables and school outcomes. The first deficiency is that the majority of empirical studies look at only a portion of the theoretical domain. Two examples are Miskal and Gerhardt (1974) who studied the relationship between perceived bureacratic structure and teacher satisfaction and Anderson and Tissier (1973) who studied the effects of school decentralization on student aspirations. Miskel et al. (1979) state that there have been few attempts to study effectiveness in any multidimensional framework where multiple criteria are related to structural or technological variables.

The second deficiency of the literature on educational structure according to Miskel et al. (1979) is the preponderance of conceptual or theoretical papers that have a total absence of empirical data to support their positions. Andes (1971) and Hanson (1975) are cited as examples of such a deficiency.

The third deficiency has to do with the unit of analysis that has been used in much of the research. Many studies use the individual as the unit of analysis in research on variables such as decentralization, formalization, bureacratic structure, and climate (e.g., Nirenberg, 1977; Odetola et al.,

1972). Miskel et al. (1979) maintain that technological and structural variables of school systems are more appropriately studied from the perspective of the school or the district. If school processes are the focus of the research, the unit of analysis should be the larger organization. Hannan, Freeman, and Meyer (1975) make a similar argument in their critique of Bidwell and Kasarda's (1975) study of school systems' effectiveness. Hannan et al. contend that the use of a unit of analysis with data aggregated above the theoretically appropriate level can lead to the "inflation of errors of mispecification" (p. 139).

To illustrate their argument, they reconstructed a data set comparable to that used by Bidwell and Kasarda to illustrate the manner in which the unit of analysis problem can lead to overestimates of the impact of organizational variables on effectiveness (as defined by achievement test scores). Their reanalysis suggests that the effects of organizational variables are smaller than that reported by Bidwell and Kasarda (1975). Such a finding is certainly not justification to exclude organizational variables from effectiveness models, but does suggest that careful attention needs to be paid to the unit of analysis, as recommended by Miskel et al. (1979). It should be pointed out, however, that aggregated data does not always produce inflated correlations as suggested by Hannan et al. (1975). Aggregation can in fact result in more unreliable data than individual data (e.g., where there is substantial variability in abilities or attitudes within a given work unit; Schneider, 1978). Thus, in the selection of a unit of analysis, the practitioner should be concerned with the interpretational compatibility between the statistical and the theoretical models selected for study.

In a well designed study relating school structural variables to multiple evaluative criteria, Miskel et al. (1979) employed the school as the unit of analysis in their investigation of the effects of several structural and process variables on multiple outcomes of perceived organizational effectiveness. Adapting the theoretical perspective of Mott (1972), Miskel et al. (1979) defined effective schools as "those schools perceived to produce products and services in greater quantity, with better quality; to show flexibility; and to exhibit adaptability to a greater extent than less effective organizations" (1979, p. 99).

The three dependent variables used in the study were loyalty, job satisfaction, and an overall measure of school effectiveness. In previous research (e.g., Hoy, Tarter, & Forsyth, 1978; Ratsoy, 1973), teacher loyalty and satisfaction had been shown to be related to group productivity. Miskel et al. (1979) adopted the position taken by Campbell et al. (1974) who believe effectiveness is best assessed by perceptions of relevant constituencies. Thus, perceptions of school system effectiveness were recorded from teachers at 125 elementary and secondary schools. Much like the functional-structures research by Pennings and his associates (1975, 1976), one independent variable of interest was organizational structure (defined as the relationship among different roles that are created to achieve educational goals. Variables such as stratification, complexity, and centralization fall in this general category of structure. In addition, Miskel et al. (1979) studied several organizational process variables in a framework compatible with Likert's (1967) organizational development model (see below). This model is concerned with factors such as leader behavior, motivational

forces, and interaction-influence patterns. Miskel et al. used Likert's (1978) Profile of a School questionnaire to assess these factors.

With results that corroborate non-educational research in a functional-structural framework (e.g., Pennings, 1975), results indicated that effective schools are perceived as having higher professional development and being more participative, decentralized, and formalized around rules. Miskel et al. (1979) conclued that environmental contingencies (e.g., community characteristics) can have a strong impact on outcome variables. Results supported research by Hoy, Newland, and Beazovsky (1977) who found that centralization reduces morale but that increased formalization increased teacher satisfaction. Findings regarding participation were also in line with predictions from Likert (1967) on System 4 managerial style (see below).

As stated above, the selection of a unit of analysis in studies of school production functions or effectiveness is critical in terms of the interpretation of the results. The issue is usually discussed in terms of the use of aggregate data to make inferences about individuals (e.g., Morgenstern, 1963; Robinson, 1950). Directly related to this issue is the assumption a researcher makes with regard to the independence/interdependence of observation in the study. Research on educational production functions, for example, has consistently assumed the independence of observations when statistical inferences are made (e.g., Hanushek, 1977; Law, 1977). Glick and Roberts (1978) discuss this issue with respect to major theories of human motivation such as expectancy and equity theory. The problem is treated in the context of bivariate analysis of individuals versus groups of individuals, holding all other important characteristics constant. They

state empirical verification must determine the degree of interdependence of the "units" in a study. Thus, in the educational context, the extent to which the observation of an achievement level of a given student is dependent on the achievement level of other students, and the extent to which a classroom's achievement level is affected by other classrooms must be empirically determined in order to properly partition the variables under study. Cronbach (1976) and Firebaugh (1978) point out that the aggregated construct is usually recognized to be separate from the individual construct with different assumptions regarding antecedents and consequences. Thus, for example, school effectiveness is not the same thing as individual performance.

Group scores (e.g., scores aggregated at the classroom, school, or district level) are considered appropriate when a group score is used to predict some individual variable. If the group score can, however, be broken down into individual scores, the use of group boundaries provides for the partitioning of within-and between-group portions of the variance. This is particularly important in breaking down classroom effects from, for example, school effects. In terms of the partitioning process, there are several suggestions (e.g., Farkas, 1974; Glick & Roberts, 1978; Kraemer, 1978). As far as predicting individual student achievement, however, (the most common dependent measure in production function and school effectiveness research), the partitioning process that makes the most sense favors interpretation of the maximum amount of variance at the individual unit of analysis (see the arguments made by Freeman, 1978 and Hannan et al., 1976).

Glick and Roberts (1978) state that a definition of group boundaries is necessary whenever there is a hypothesis as to the interdependence of a

set of individual scores. The extent to which regression and correlation coefficients are biased can only be determined if estimates of "compositional effects" (see Werts & Linn, 1971) can be made. Unless these contextual variables can be estimated, it is impossible to make inferences across levels of analysis as, for example, in Bidwell and Kasarda (1975). In assessing the relative contributions of school-related variables on student outcomes, the group boundaries must be specified, the extent of their interdependence estimated, and, if possible, the direction of causality hypothesized between groups.

In addition to the concerns expressed above, several authors (e.g., Bidwell, 1965; Mohr, 1973; Pellegrin, 1976) have commented on the lack of good instrumentation in studies of school structural and technological variables. One example of good measurement is found in Bishop and George (1975) who developed a "Structural Properties Questionnaire" to measure an organization's characteristics. This questionnaire has been used as the measure of school bureaucracy in several studies (Miskel et al., 1979; Murphy, Bishop, & George, 1975). However, for the most part, research in education that most clearly matches the functional-structures model of effectiveness is plagued by the same problems of design and methodology discussed by Pennings (1975). While the functional-structures model itself looks promising as a model of school system effectiveness, more research, utilizing multivariate designs and larger units of analysis (e.g., school or district), such as the research conducted by Miskel et al. (1979), is needed before a definitive statement can be made on the appropriateness of the functional-structures model for school systems.

The Managerial Process Model

The managerial process model operationalizes an organization's effectiveness within the context of abilities to perform certain managerial functions. For example, the extent to which a principal or school administrator effectively performs the functions of decision making, budgeting, or planning would be a common research strategy. As denoted by its title, the focus of this model is on the administrator's, or manager's, effectiveness, particularly in integrating individual goals with organizational goals. Thus, measures of employee job satisfaction, motivation, productivity, and beliefs are the most likely dependent measures for this model. Some discussions of this model sound very much like discussions of Likert's "principle of supportive relationships" (Likert, 1961). Predictors of managerial success in decision-making, etc. through motivational or attitudinal measures are likely independent variables of interest.

Reiman (1975) studied managerial attitudes toward "critical others" who could potentially affect an organization. This research is one of the better representations of the managerial process model. His "Public Values of Management" questionnaire was developed to study the relationship between public-sector managerial attitudes and effectiveness as measured by performance. Scores on the questionnaire were found to be significantly related to group performance measures. While several writers in the educational domain make reference to the crucial nature of the principal's functions and consider them in the context of managerial theory, there have been few published attempts to empirically study the relationship between attitudes of principals and educational outcomes (e.g., Barro, 1970; Dyer, 1970; Heil,

1962; Radnor, 1974; Silberman, 1970; Stanchfield, 1976). Empirical research in the educational setting has concentrated on variables such as principal race, age, sex, educational attainment, tenure, salary, and full or part-time positions as they relate to various outcome measures (e.g., student gain scores; Averch et al., 1972; Coleman et al., 1972; Hanushek, 1977). In general, factors related to the principal have been found to contribute fairly trivial amounts of variance in the outcome measures. More "job-related" variables within both the ability and the attitudinal domain would seem to be a richer source for predicting outcome variables. For example, Mahoney and Weitzel (1969) concluded that effectiveness was a result of the additive combination of productivity-orientation, planning, reliability and initiative as related to managerial functions. Reliably measured ability and attitudinal variables for principals that are directed at these issues would seem to have a far greater potential for predicting educational outcomes.

Other researchers have looked at school administrators in their managerial capacities and discussed problems in administrator-teacher relationships and their effects on effectiveness (e.g., Cicourel, 1975; Corwin, 1970; Dornbush & Scott, 1975). Corwin (1970), for example, found that perceptions of a lack of administrative backing, conflict over students, and lack of authority resulted in lower teacher morale. No attempt was made to relate this to student output measures, however. Bidwell and Kasarda (1975) found that administrative intensity depressed median levels of achievement in their study of 104 Colorado school districts. Their explanation was that administrative intensity (as measured by staff-teacher ratios) merely diverts human

resources from teaching and other staff functions related to teaching. Their results have never been replicated and the practical significance of the administrative intensity effect has been questioned (Alexander & Griffin, 1975).

Barro (1970) probably comes the closest to a conceptualization of school system effectiveness in terms of a managerial process model. He recommends that a planning-programming-budgeting method of management be implemented that is more output-oriented than present systems. This approach is presented in an accountability framework such as that presented by Dyer (1970). Similarly, Radnor (1974) implicitly endorses a more systems-oriented, input-output model to replace the pervasive "management by exception" style.

In other discussions of educational management (e.g., Yee, 1973), systems management models are presented which, while somewhat similar to the managerial process models (e.g., Cyert & March, 1963), appear to be a closer match with the systems-resource model, discussed above (e.g., Georgopoulos & Tannenbaum, 1957). The managerial process model is, however, more focused on investigating the characteristics, behaviors, and effectiveness of administrators.

There is an increasing interest in the importance of educational management on traditional output measures, as evidenced by the attempts to improve selection procedures for school administrators and principals. For example, the National Association of Secondary school Principals has developed a comprehensive assessment center to aid in the selection of principals and assistant principals (Hersey, 1977). The program attempts to identify performance in the assessment center that is related to high performance as administrators. A distinction, however, must be drawn at this point between managerial behavior,

performance, and effectiveness. Campbell, Dunnette, Lawler and Weick (1970) state that behavior is simply what people do in the course of working (e.g., dictating a letter, or directing a subordinate, etc.), while performance is behavior that has been evaluated (measured) in terms of its contribution to the goals of the organization. Finally, effectiveness refers to some overall index of organizational outcomes for which an individual is at least partially responsible, such as achievement test scores, teacher satisfaction, turnover, etc.

Since the measure of managerial effectiveness is one or more steps removed from what the individual actually does, managerial performance is likely to be more easily predictable than managerial effectiveness, as performance depends directly upon managerial behavior and is not a function of additional factors not under the control of the individual (e.g., socioeconomic level of students). Preliminary results from the NASSP assessment center research supports this position. Therefore, it should be pointed out that a number of criterion measures that have been used in the past to predict school administrator performance have proved to be insensitive to actual measured effectiveness due to situational or moderating variables that can affect effectiveness (e.g., Coleman, 1972). Additionally, the manager's job is multifaceted (Ghiselli, 1956) such that different managerial behaviors may result in the same managerial performance rating. Also, a particular managerial behavior may be evaluated as being highly effective according to one criterion, yet ineffective using another standard. See, for example, Beshers' (1972) discussion of conflicting administrative goals in education. The question we need to ask in terms of the educator-manager,

as discussed by Campbell et al. (1970), is:

> What are the varieties or combinations of organizational circumstances, personal characteristics, and behavior patterns that are likely to be perceived as effective managing?

Predictably, the next question would be: what is it that is perceived as being "effective" management in education? England (1966, 1967) in surveying more than 1000 managers throughout the country, determined the goals that were perceived as being both important and highly indicative of successful managing. Topping the list were the objectives of organizational efficiency, high productivity and profit maximization. Therefore, it is reasonable to define effective management in terms of organizational outcomes (Seashore & Yuchtman, 1967):

> We define effective managerial job behavior as any set of managerial actions believed to be optimal for identifying, assimilating and utilizing both internal and external resources toward sustaining, over the long term, functioning of the organizational unit for which a manager has some degree of responsibility (p. 344).

Effective managers, then, optimize material, human, and financial resources in sustaining the operation of the organization. This operational definition relieves us of the evaluation of traits that are generally used to describe the successful or unsuccessful administrator-manager and allows us to concentrate on the tasks at hand. Unfortunately, in the educational context there is little empirical support for a principal performance-educational outcome relationship.

Stewart (1967) lists three methods commonly used to study what managers do on their jobs: An observer may record what a manager does, the manager may keep his own records or diary of his activities during a workday, or the manager may be asked merely to estimate how he spends his time, using a prepared checklist of job duties or behaviors. Numerous studies have utilized these approaches, with varying but compatible results (Burns, 1957; Carlson, 1951; Horne & Lupton, 1965). One of the earliest job description checklists used to delineate the responsibilities of managerial personnel was reported by Guion (1953). Cluster analysis resulted in five areas of responsibility, namely: 1) labor relations; 2) methods and rates; 3) maintenance of records; 4) personnel selection and placement; and 5) meeting and coping with day to day operational problems as they arise. These results closely match the results of job analysis of school administrators (e.g., Hersey, 1977; Yee, 1975).

Another approach to description involves behavior sampling in which the type of activity involved in at random intervals during the day is recorded. While this approach has been used to infer the determinants of managerial effectiveness (Kelly, 1964), O'Neil and Kubany (1959) were able to conclude from their research that the mechanistic recording of job activities is incapable of revealing crucial differences between more and less effective foremen. Thus, for the practitioner, the utility of this approach in measuring effective administrative performance is questionable.

The critical incident method developed by Flanagan (1954) is, however, a more promising technique for sampling administrative behavior. This approach requires reports by qualified observers of the things that managers

do that are especially effective or ineffective in accomplishing parts of jobs. After a number of such behavioral incidents are collected, the abstracted categorized incidents are combined to form a composite picture of job essentials. With continuous recording and reporting of critical incidents, this technique can reveal the dynamic attributes of a particular managerial job, and since it focuses on managerial performance, i.e. what the manager actually does, can be expected to yield behavioral information uncomplicated by situational variables. Several investigators (Anderson & Nilsson, 1964; Kay, 1959; Williams, 1956) have used this to advantage, demonstrating a high degree of inter-rater reliability in developing nonambiguous category definitions, based on critical incidents. In seeking to establish the reliability of incident classification and organization, Smith and Kendall (1963) suggested that behavioral incidents be retranslated by those actually using the scale to determine those incidents that can be reliably assigned to a particular dimensional category, while eliminating statements and dimensions that fail to delineate highly specific and non-ambiguous job behaviors that can be reliably used to develop a final set of job specific behavioral scales. Bernardin and Smith (1981) have argued the approach has potential for alleviating common errors in performance measurement.

In the educational literature no studies could be found which used critical incident methodology, retranslation, or detailed observational procedures to identify critical task components for the educational administrator. While these approaches are certainly improvements over any trait-oriented measurement of managerial individual differences, there are no guarantees that such refined measurement strategies will lead to the

establishment of a stronger relationship between educator-management performance and outcome measures such as achievement test data. Certainly we can predict a better chance for significant relationships between effectiveness and administrator or principal managerial parameters as measured above than with characteristics such as race, age, tenure, and salary that have been used most frequently in the past.

Gross and Herriott (1965) developed the "Executive Professional Leadership" (EPL) questionnaire for elementary school principals. While the measure was not based on any of the methodologies discussed above in describing the manager-principals' job, scores on the EPL were found to be related to teacher morale, teacher performance ratings, and student learning. Hilfiker (1969) found that scores on the EPL correlated positively with perceived levels of social support and system innovativeness. Teachers were asked on the EPL to indicate the extent to which their principals engaged in activities such as: "makes teacher's meeting a valuable educational activity"; "treats teachers as professional workers"; "has constructive suggestions to offer teachers in dealing with their problems."

Another measure that attempts to assess the behavior of a principal is the Profile of a School (POS; Likert, 1978, see below). In research with the POS, one of the factors derived in factor analysis was labelled "principal leadership" by Miskel et al. (1979). The authors stated that this factor described the principal's behavior in terms of supportiveness, work facilitation, goal-emphasis, and interaction facilitation. The focus of the factors was on the principal-teacher level of interaction. For example, one of the items was "how often do you see your principal's behavior as friendly and

supportive?" Results from the Miskel et al. study indicated "principal leadership" was the best predictor of teacher loyalty to the organization and teacher satisfaction.

The Leader Behavior Description Questionnaire (LBDQ) is an established method for describing the behavior of leaders. While there are no published studies relating LBDQ scores of principals to effectiveness measures, Halpin (1956) used the LBDQ in a study of school superintendents. Respondents were asked to complete the LBDQ in two ways: first, according to their perceptions of the superintendent's actual behavior and second, based on their expectations on how the superintendents should behave. Halpin found large discrepancies between the actual and the ideal perceptions of the superintendents' behavior. Such discrepancies were found to be negatively correlated with teacher satisfaction. Similar results have been found in other research relating school superintendent behavior to teacher satisfaction (e.g., Bidwell, 1965; ERIC, 1973; Guba & Bidwell, 1957).

It is apparent that managerial process variables for school administrators have been shown to be reliably related to several of the "soft" criteria used to assess organizational effectiveness (e.g., teacher satisfaction and loyalty). However, few studies have established a link between managerial behavior and any "hard" criteria such as efficiency or student outcomes. Therefore, for the present, the practitioner is cautioned not to use managerial process variables as direct and unequivocal indices of school system effectiveness. If pursued, research into the relationship between managerial behavior and performance in schools should shed some light as to the utlity of this potentially fruitful approach.

The Organizational Development Model

The Organizational Development (OD) model views effectiveness in terms of the organization's problem-solving and renewal capabilities (Cunningham, 1977). Utilizing the knowledge and intervention techniques of the behavioral sciences, this model attempts to shape the organizational structure and climate such that the individual's ability to satisfy growth needs is maximized. It is only when the potential of each member is fully realized that the organization can be maximally effective.

OD has been called a normative model because it attempts to prescribe the requisite conditions under which an organization will be effective (Steers, 1976). These conditions would include: 1) being aware of, open to, and reactive to change: 2) searching for new methods and forms of organizing; 3) having an optimistic view of its members - self-actualizing and trusting; 4) seeking to insure the satisfaction of its members since they are the reason for the organization's existence; 5) having open communication; and 6) confronting and dealing with conflict directly without personal references.

OD, as can be seen, has to do with the human component of the organization. It shares many features with the managerial process model just described. Both models examine the behaviors of individuals in organizations, and both models focus on the individual's perceptions of and reactions to the managerial processes operating. Furthermore, both assume that improvement in the informal organization will result in a more effective organization (Cunningham, 1977). However, unlike other system approaches, OD seldom mentions goal outcomes. If such things as profit and turnover are mentioned

at all, it is in a fairly unsystematic way and only after much discussion about such employee-oriented factors as increased individual openness, better communication, grater individual self-actualization, and other indicators of what is considered to be a healthy system (Campbell, 1976) from the individual's perspective. No mention, however, is made of the students', parents', or community's perspective.

The OD model only assumes that if an organization (e.g., school system) can achieve the list of end states that define a healthy system (e.g., Beckhard, 1969), it will be effective as an organization and will be optimally equipped to carry out its mission(s). However, the model makes no mention of what the mission is, or how the organization attempts to achieve it.

Typical of an OD model is the attempt to prescribe the requisite conditions under which an organization will be effective, without mention of a rational or empirical defense of the criteria as true measures of effectiveness. The OD model assumes organizational effectiveness without ever measuring it.

The OD model is related conceptually and empirically to the literature on organizational climate in this respect. Organizational climate has been described as the result of the transaction between individual members, with their individual needs, abilities and goals, and the organizational structure (Campbell et al., 1974; Schneider, 1980). It reflects the strength of the prevalent values, norms, behaviors and attitudes of the organization. Organizational climate has been studied rather extensively in the field of education. Numerous climate measures have been developed for educational environ-

ments, the most frequently used being Stern's College Characteristics Index, Pace's College and University Environment Scales, Aston and Holland's Environmental Assessment Technique, Pervin's Transactional Analysis of Personality and Environment, and Halpin and Croft's Organizational Climate Description Questionnaire (Campbell et al., 1974). The Organizational Climate Description Questionnaire (OCDQ) (Halpin & Croft, 1962) is one of the few climate measures that has looked specifically at the climate of elementary schools. This instrument measures responses to descriptive statements about peer (teacher) and principal behavior. Four of its eight scales tap teacher perceptions of typical peer group behavior and the remaining four scales assess perceptions of the behavior of principals in their interactions with teachers. Results from the OCDQ are interpreted in reference to an "Openness vs. Controlled" continuum. Theoretically, any school can be placed on such a continuum depending on its characteristic pattern of internal social behaviors. The practical importance of climate research in education has suffered however, because in general it has not attempted to measure school system effectiveness. Much like the OD model, it only assumes effectiveness. Furthermore, there has been virtually no attempt to relate the limited research on school climates to a vast research on organizational climate (Schneider, 1980).

Radnor (1974) has discussed early applications of OD in education and finds that, for the most part, schools do not perceive any need for improvement in organizational functioning. Numerous examples of ill-fated, OD interventions can be cited which started in a climate where participants did not perceive a great need for change. These studies most often use employee

attitudinal measures to assess change as a function of the intervention. Thus, not only is the OD approach plagued by measurement problems in terms of objectives but also in terms of perceived needs for such an approach. One study that fits into the classical, OD framework found a significant relationship between perceptions of "openness" in the school and job satisfaction of the teachers (DiCaprio, 1974). Such data serves as the basis of some type of OD intervention to facilitate greater "openness" in certain schools. However, as stated earlier, there is a paucity of data that relates teacher attitudinal measures such as job satisfaction to any effectiveness measures such as student achievement, teacher turnover, absenteeism, etc.

While there are differences in specific characteristics of organizations that OD theorists define as "healthy" systems, there are three general assumptions underlying most OD approaches. The most important assumption concerns the general nature of man and is best represented by McGregor's Theory Y (McGregor, 1960). This philosophy of management states that the expenditure of physical and mental effort is as natural as play or rest and that work is a source of high satisfaction. Also, people will exercise self-control and self-direction in the service of objectives to which they are committed. The second assumption is that organizations must anticipate and cope with a changing world. Thus, a rigid, functional bureaucracy is outmoded in this dynamic environment (Bennis, 1969). The final assumption implicit in the literature on OD is that organizations exist primarily for the benefit of members of the organizations (Campbell et al., 1974). One example of an operationalization of these assumptions is the work of Staunfield (1976).

He studied the effects of teacher-administrator interactions on elementary reading achievement. His major hypothesis was that the greater the participation by the administrator in program planning and "supporting" teachers, the greater the outcomes. The approach is a close match to French's (1972) description of a "healthy" system, characterized by high trust and support across organizational levels, open communication, group responsibility for design and planning, confrontation rather than avoidance of problems, and frequent synergistic solutions. Staunfield found that such a climate was in fact significantly related to reading achievement.

Another example of an OD approach to education is Mardirosian (1975) who reflects the theoretical work of Bennis (1969). Bennis' characteristics for an effective organization are: decision-making located close to the sources of information; a reward system linked to the achievement of organizational goals; and a support system built into the vertical, organizational hierarchy. Mardirosian found that teachers from higher achieving schools in the state of California perceived stronger support systems and closer relationships between rewards and goal attainment than teachers from the low-achievement schools.

A final example of a traditional OD orientation is the work of Corwin (1970), in a study of educational effectiveness. His study of organizational conflict in high schools is a vivid description of an "unhealthy" system as defined by Beckhard (1969): conflicts avoided or supressed; closed communication channels; no concern for subordinate growth or development; and no formal feedback systems for performance. Teachers in Corwin's study indicated that factors such as peer incompetence, lack of supervisory

support, and constricted authority had deleterious effects on their performance. As with most studies conducted in an OD perspective, however, no actual performance or effectiveness data were collected.

The major shortcoming, then, of the OD models discussed above as applied to school systems is that they either assume or ignore a linkage between OD goals and the more clearly defined, objective goals of school effectiveness such as student performance. These latter goals are the very reason for the school system's existence, not the psychological "growth" of organizational members such as teachers, administrators, etc. A useful model of school system effectiveness must go beyond an assumption that school systems are effective solely on the basis of teachers having "self-actualized", developed better communication skills, or higher levels of satisfaction. The model of effectiveness must include both clearly stated objectives and the criteria to be used in measuring these objectives. Models of organizational development include neither of these dimensions and are therefore not adequate as models of school system effectiveness.

The Likert-ISR Model

The Likert-ISR model of organizational effectiveness (e.g., Likert, 1967) can be classified as an OD model in that its major focus is the notion of shared power or participation in decision-making. Thus, there is overlap between the Likert-ISR model and the OD models promulgated by Bennis (1966), French (1972), Shein (1965), and Beckhard (1969). The major premise of the Likert-ISR model is that the closer the administration of a school system is to "System 4", the more effective the system is viewed to be. System 4 is a democratic, participative group model. All decisions are reached by group

consensus, and all groups have "linking pins" (i.e., persons who hold overlapping membership in groups). It is assumed that since decisions are made by group consensus, all decisions will be fully accepted both overtly and covertly. Communication lines in System 4 are open and move freely among and between all levels of the organization, with little or no distortion or filtering of that information. Because fear or coercion are not used, attitudes are very positive. A climate of trust and openness is the rule, not the exception.

In order to assess the extent to which a school system is participative in managerial style, Likert (1978) developed the Profile of a School (POS). The POS consists of 10 questionnaires that are designed to record behavioral information within the organization as seen from the various levels of the organization (e.g., students, teachers, principal, superintendent, and parent). While the POS attempts to characterize individual's overall attitude and motivation, its primary concern is to record current behavior and organizational practices at the various levels of the organization. The variables it attempts to measure are those that have been found to be related to quality of performance in both school administration and business settings: leadership, decision making, problem solving, motivation, communication, conflict management, interaction, and the structure through which interaction occurs (Likert, 1978). With the POS, schools or school districts can be diagnosed as either System 1 (authoritarian, non-participate), System 2 (benevolent authoritarian), System 3 (consultative) or System 4 (fully participative).

Field research comparing System 4 schools with their counterpart (System 1) has been summarized in Table 3.

Insert Table 3 about here

A few of these studies have correlated objective data such as achievement test scores with system classification and have found both higher achievement test scores for System 4 (Donlan, 1978; Gibson, 1974) and no relationship between managerial style and educational performance (Horsman, 1973). There is, however, a considerable body of field research that does support the validity of the Likert-ISR model, as shown in Table 3. One limitation of most of these studies is that the outcome measures have been "soft" ones, such as job satisfaction, need satisfaction, labor relations, morale, etc. The limitations of such "soft" criteria have been discussed previously.

The Likert-ISR model also fails like the other OD models in that it assumes that if participative managerial conditions exist, the school system will be effective. Stated differently, the Likert-ISR model proposes, in essence, that all a school system need do to have an effective organization is to have a System 4 organization as they define it. While this reasoning is questionable in and of itself, Likert and his associates support this logic with research that often uses only one or two of the 10 POS scales developed. Thus, the research purporting to support the Likert-ISR model often only gathers information at one or two levels of analysis (e.g., the teacher and the principal) and yet implications are made in terms of the entire school system. Despite these limitations in the research, the Likert-ISR model itself does have much to offer in terms of assessing the perceived effectiveness of a school system. The POS can be used as a

diagnostic tool to assess the internal state of the organization, and the Likert model can suggest ways in which to improve that internal organizational state. However, since organizational climate is not the only determinant of organizational effectiveness, the Likert model is not sufficient in and of itself to assess organizational effectiveness.

Conclusion - The Choice of a Model for School Systems

It is our contention that what the literature on organizational effectiveness needs least is another model of effectiveness. Rather, we believe there are components of several of the models discussed above that are appropriate for school systems. The advantages and disadvantages of each of the models has been delineated above.

The two models most supported by empirical research are the management-by-objectives and the Likert-ISR models. Both models have problems when applied to school systems. We believe the positive features of these models, plus those of the other effectiveness models are best conceptualized in a process model of effectiveness.

Campbell et al. (1974) and Steers (1977), in the general context of organizational effectiveness, and Madaus et al. (1980), in the context of school effectiveness, have recommended models of effectiveness with emphasis on processes related to effectiveness. In all three cases, the process models incorporate various features of goal-centered and systems-oriented models of effectiveness. Steers (1977), for example, delineates three major dimensions for his process model. These dimensions are: 1) goal optimization; 2) a systems perspective; and 3) a behavioral emphasis. In reference to many of the goal-oriented approaches to effectiveness, he states assessments

of an organization's successes or failures must be made with reference to the intentions for the organization rather than the value judgments of the researchers. However, Madaus et al. (1980) point out that such intentions (i.e., school objectives) are indefatigably linked to the value judgments of those responsible for the school. Thus, only if some consensual agreement can be reached with regard to the "intentions" for schools can absolute or relative comparisons be made with regard to school effectiveness. While Madaus et al. (1980), Bloom (1976) and many others agree that the process of reaching such consensual agreement will indeed be difficult, the majority also feel the task is not insurmountable. Campbell et al. (1974) present an excellent procedure for the development of task objectives based on the perspectives of all relevant constituencies. The adoption of this procedure, a varient of the scaled expectancies methodology from Smith and Kendall (1963), would provide for an operational definition of "consensual agreement" of intentions and a documentation of the differences in intentions as a function of the constituency. Likert's (1978) "Profile of a School" questionnaire has items related to intentions from the perspectives of board members, principals, teachers, and students.

Steers (1977) also strongly endorses the open systems theory of organizational effectiveness which emphasizes the interaction of the organization with the environment. The major focus of this perspective is on the relationships between components both inside and outside the organizations as they affect the organization. The discussion by Metfessel and Michael (1967) in the context of colleges and universities is a good example of an open system, goal-optimization model which also has applications for elementary

and secondary school systems. Program evaluation, they state, should involve members of the total school community such as lay individuals and groups, professional personnel, and students. Likewise, in incorporating an open-systems perspective for school effectiveness, relationships between constituencies representing the total school communities should also be recognized and, according to Steers (1977), should make it easier for managers to "... take decisive actions to facilitate goal attainment because of their increased understanding of organizational dynamics" (1977, p. 177). Thus, Steer's thinking reflects one of the major components of the managerial process model.

The third dimension of Steer's (1977) process model emphasizes the role of behavior in affecting organizational performance. Steer's states that if we are to gain a better understanding of the factors that explain effectiveness, it is critical to first gain an understanding of the basic unit of analysis; that is, the behavior of the organization's members. It is this third dimension that is most often missing in the goal-oriented or systems-resource models. As stated in the reviews of the specific models within the goal-oriented and systems perspectives, there is generally little consideration given to the "means" (i.e., the behaviors) employed for organizational "end-states". While this criticism is more appropriate for the goal-oriented models (i.e., rational-goal, MBO, functional, and cost-benefit), the systems models are also generally outcome-oriented, particularly the systems-resource and functional-structures models. In fact, only the managerial process model considers the "means" (i.e., managerial behavior) as an important element of the model. Unfortunately, as discussed

above, there are few empirical studies of managerial behavior in school systems.

The greatest gap in the research on school effectiveness is in this general area of human resource behavior and performance. There are, of course, a great many studies which focus on teacher behavior and its relation to certain student variables. However, none of these studies consider these relationships in the context of school effectiveness. There is a need to assess such input and output variables within the framework of a process-oriented, behaviorally-based model of effectiveness. Relationships between task objectives at the macro level, organizational member behavior at the micro level, and outcome measures such as efficiency (systems research), climate (functional-structures), morale (Likert-ISR) and student outcomes (MBO) should be investigated.

There has been very little research which has attempted to employ a multivariate strategy in order to study how the various school outcome measures interact. Also, with the exception of the studies on compensatory school programs, the use of static research designs has precluded investigations of the dynamic nature of the various criteria employed in studies of school effectiveness and their relationships to various school inputs. The use of archival data in order to make hypotheses about these relationships would certainly be an excellent first step in this regard.

Human Resource Process Variables in the Process Model. There was generally a lack of research in the area of the human resources available to school systems (i.e., teachers, principals, administrators) and human resource policy. Studies that have investigated human resource variables

have concentrated on those that are most easily retrievable such as years of education or experience, salary, and demographic background characteristics. A process model of school effectiveness should incorporate factors pertaining to personnel selection, training, and evaluation procedures, particularly for teachers and principals. Based on a considerable amount of research in these areas outside of education, we can hypothesis that some "process" variables in school administration are related to uncontaminated measures of school effectiveness. For example, the extent to which a content-domain sampled, structured interview is used with trained, multiple assessors should be related to more valid selection decisions at both the teacher and principal levels (Dunnette & Borman, 1979). The type of formal performance feedback system established for teachers should be related to improvement of teacher performance over time. The level of teacher job satisfaction should be related to both absenteeism and, to a lesser extent, turnover. The level of professionalism exemplified in schools by ongoing training programs and teacher-principal interactions should be related to improved teacher performance.

We believe the step following the task objectives procedure described by Campbell et al. (1974) in the refinement of a process model of school effectiveness should be the development of a questionnaire that assesses variables related to school personnel selection, training, and evaluation policy. Such a questionnaire should also include organizational climate variables shown by Likert (1978) and others to be related to several outcome measures. As will be discussed below, we believe the personnel process variables will show relatively little variance across school systems

at the present time.

Madaus et al. (1980) discuss several other factors found to be related to higher student outcomes in several studies. While there was considerable overlap with some of the climate variables discussed by Likert (1978), two important predictive variables were student perceptions of academic futility and teacher "push" (e.g., Brookover, Schweitzer, Schneider, Beady, Flood & Wisenbaker, 1978). Brookover et al. (1978) reported that the social-psychological and normative variables, based on these measures of climate, contributed close to 80% of the variance in mean school achievement when such variables are entered first in a regression analysis. Madaus et al. (1980) also cite several other studies from their own work outside of the United States to support the importance of such climate variables. Items related to these variables would thus be appropriate for the "process" questionnaire.

This "process" questionnaire should be based on the results of the task objectives procedure and incorporate the salient features of the goal-oriented and systems-resource models discussed above. Efforts should be made to theoretically link these process variables to outcome measures frequently used in the educational literature. Before an empirical study relating the "process" responses to outcome measures is undertaken, however, major considerations must be given to the important issues of research design and data aggregation. The next section will critique research designs that have been employed to investigate variables related to school effectiveness.

Research Designs in Studies of School Effectiveness

While there is near unanimous agreement that, among school related variables, teaching quality is the best predictor of student outcomes, no studies could be found with an appropriate design for investigating the effects of teachers over a period of time longer than one year. Several studies, discussed above, have found moderate relationships between teacher characteristics and measures of student achievement (e.g., Rosenshine & Furst, 1971). However, attempts to study the effects of schooling at a higher level of aggregation (e.g., school, district, or state level) have made no provisions for the fact that individual teacher effects will be diluted in any such analysis. Consider, for example, the following research scenario: pupils in grades 3, 6, and 9 are tested for achievement and their scores are aggregated over teachers and schools to the system level. The scores are then correlated with a great number of potential predictors of student achievement. The results reveal trivial effects for variables related to the school, including those of the classroom teacher. Based on this type of research, the conclusion is drawn that "school brings little to bear upon a child's achievement that is independent of his background and general social context" (Coleman et al., 1966, p. 325) or "schooling does relatively little of what it always claimed to do - foster the cognitive abilities of students and thus assure them a productive niche in the economy" (Levine & Bane, 1975, p. 19). While Madaus et al. (1980) have done an excellent job of reviewing the problems in this type of design and the inferences drawn from research with it, the impact of the data aggregation method on teacher effects should be emphasized. Roberts, Hulin, and Rousseau (1978) state

that if one is going to aggregate data prior to data analysis, the interpretation of the meaning of the data should be clear prior to aggregation. Given the dire conclusions drawn from aggregated data analysis such as in Coleman et al. (1966) and Jencks et al. (1972), it is clear there was little consideration of interpretation vis a' vis aggregation.

The ideal study to investigate teacher effects would be something like the following: 1) using a valid performance assessment method, teachers are rank ordered or rated on several important dimensions of effectiveness within grade levels; 2) criteria are selected that have construct validity for school effectiveness and conceptual similarity with the dimensions of teacher effectiveness (e.g., student achievement and teacher "achievement-orientation" or student motivation and teacher "support facilitation"); 3) a large group of students ($n \geq 600$) is randomly assigned or matched (based on background characteristics) to teachers who have been identified as "highly effective" on the dimensions of effectiveness or to teachers identified as "less than highly effective" (the ideal design would identify teachers of both high and low effectiveness on the dimensions). Assuming we can identify teacher groups of both high and low effectiveness on two dimensions, our longitudinal design could be something like a 2 x 2 x 6 (grade levels) where we would study individual students over a six year period; 4) students are assessed on the criteria over grade levels with greater effects predicted over time as a function of teacher effectiveness.

Before we are locked up and put away, let us emphasize that this "ideal" design has been presented for illustrative purposes only. We only offer it as a basis for comparison to the designs that have been used in the past in

order to "tease out" the effects of teacher performance or other school-related variables that do not affect all students equally. Based on the four steps in the "ideal" design, let us make some general comments about the quality of the school effects research we have reviewed:

1. Performance assessment procedures for teacher effects are crude. For the most part, criteria used to operationalize teacher performance has been primitive with little consideration of research implications in the area of performance assessment. The majority of studies reviewed used trait-based, supervisory ratings of teacher performance, a method replete with potential for error (Bloom, 1976). The much fewer number of studies that investigated principal or administrator performance are also generally guilty of faulty methodology in this same respect.

Other studies have employed measurement strategies with even less potential for validity. Coleman et al. (1966), for example, used teacher verbal ability as a predictor of student aptitude. While there was some evidence that teachers' verbal ability was related to achievement, the variance accounted for by this predictor was low. This does not strike us as surprising. An analogy in the measurement of managerial performance would be to use height as the measure of success. Several studies have documented the correlation between height and managerial performance so why not just use height as the measure of success rather than bothering with the cumbersome task of measuring performance. The simple answer is that height is <u>not</u> a measure of performance and neither is the verbal ability of teachers. Both variables are only correlates of the constructs under study and since their respective correlations with performance in no way approach those of acceptable reliability coefficients,

they should not serve as replacements for the variables we wish to study. Of course, given the sample size (645,000 students and 60,000 teachers) and time constraints for the Coleman et al. (1966) study, it is understandable why a verbal ability measure was used as a surrogate measure of teaching quality. However, it is equally understandable why such teacher effects contributed so little to the predictive equation. The major point of this argument is that if your interest is teacher effects, efforts should be made to maximize the predictive potential of these effects with sound measurement strategies (cf., Rosenshine & Furst, 1971).

2. The most common criterion for the assessment of school effectiveness has been scores on standardized tests (Averch et al., 1972; Madaus et al., 1980). Cooley and Lohnes (1976) have taken the position that general verbal and nonverbal ability tests are the preferable measures of school effectiveness for both primary and secondary levels because the "intellectual factor" measured on such tests is the most common across schools and grade levels. Coleman et al. (1966) also argued for the preference of verbal and non-verbal ability tests.

The issue of the appropriateness of various outcome measures for studies in school effectiveness is best conceptualized as an issue of measurement validity. Validity in the context of school effectiveness is the extent to which outcome measures correspond to actual effectiveness levels for school systems. Just as when we are interpreting a test we wish to make an <u>inference</u> about the degree to which a person possesses a certain trait or quality measured by the test so also in interpreting effectiveness data, we wish to infer that a school system's actual level of effectiveness on a dimension is

reflected by the effectiveness data.

The traditional types of validity are mentioned in virtually every text in educational testing and differential psychology and are discussed in detail in the Uniform Guidelines on Employee Selection Procedures, the Standards for Educational and Psychological Tests (American Psychological Association, 1974), and the Principles for the Validation and Use of Personnel Selection Procedures, Second Edition (APA, Division 14, 1980). In terms of school "effectiveness", "construct validity" is the most appropriate for the definition of validity given above (James, 1973). There can be no question, however, that features of content and criterion-related validity are also applicable. In fact, we are of the position that the distinctions that have been drawn between the different types of validities result in an oversimplification of the meaning of validity, leading to numerous reports of "one-shot" validation processes (Dunnette & Borman, 1979). The APA Standards state that "a thorough understanding of validity may require many investigations" (APA Standards, 1974, p. 25). This statement probably best sums up our position on the validity of school effectiveness data. The term "construct validity" probably best captures this position (Campbell, 1976).

Construct validity has been defined as the degree to which variability in a measure is a function of variability in some underlying construct. Thus, in the school effectiveness context, variability in the measures of school effectiveness should be a function of the variability of the actual effectiveness levels of those schools. The same definition of construct validity can be applied to teacher effectiveness by simply replacing the

word "school" with the word "teacher". What should be underscored in the definition of validity is the notion of inference. Validity does not refer to a specific measurement strategy. For example, it should never be said that MBO is a valid technique for measuring school effectiveness, implying that validity is a constant across users, uses, and situations. Rather, the inferences made from the use of MBO data can be said to be valid or invalid based on the evidence presented. Cronbach (1971) put it best when he said " ... one validates not a test, but an interpretation of data arising from a specified procedure ... Because every interpretation has its own degree of validity, one can never reach the simple conclusion that a particular test is valid."

The evidence presented to support inferences of validity can and should be from many different sources and based on methodologies that have been recommended for content, criterion-related, and construct validity. It is thus important to conduct criterion-related validity studies where possible, and to use content-oriented domain sampling in order to be able to justify a conclusion that effectiveness data are valid.

Research on school effectiveness has in general ignored the issue of content-domain sampling and concentrated instead on criterion-related validity studies using ability tests scores as the criterion. Madaus et al. (1980) make convincing arguments that the generally poor criterion-related validities are at least partially a function of the lack of construct validity for the ability tests. They state that "... schools have a whole host of objectives in the cognitive area which are not represented in a test of general ability ... It seems curious to look for differences between schools on something

that is only incidentally taught while ignoring the possible differential success on material that is more specific to school curricula" (1980, p. 124).

3. Step three in the "ideal" study presented above had to do with experimental design and sample size. The "ideal" design was presented that would be most sensitive to teacher effects. Our review of the literature on school effectiveness revealed no studies employing designs that approximated the "ideal" and, with a few notable exceptions, sample sizes that rendered most research designs virtually "powerless" to refute the null hypothesis. On the issue of sample sizes, a classic article by Schmidt, Hunter, and Urry (1976) documented the statistical power of test validation research given certain sample sizes. They were able to show that given the common problems of range restriction in the criterion and criterion unreliability, most researchers grossly underestimated the number of subjects needed to detect true differences. Such a problem may also have accounted for many of the dismal results for school effects reported in the literature. With regard to the issue of research design, numerous examples of quasi-experiments (Campbell & Stanley, 1966) were found in assessments of compensatory programs such as Head Start (e.g., Cicirelli, 1970). However, no examples of quasi-experiments that approximate the "ideal" design above could be found for non-compensatory schooling. For example, it would be possible to compile criteria data on teachers and, based on ranking or ratings of performance, group students based on the number of "highly effective" versus "less than highly effective" teachers they were assigned through a designated grade level. This quasi-design would resemble the straight "ideal" experiment with the exception that subjects would be "randomly" assigned to

a number of highly effective and less effective teachers over the six year period. Predictions could then be made based on the number of "highly effective" versus "less than highly effective" teachers (between dimensions) and the effects of different distributions of teaching quality over the six year period. For example, the effects of where in the six year period the students received the "highly effective teaching" could be investigated.

While this quasi-experimental design would not be as sensitive as the "ideal" design presented above, assuming there is reliable variance in the predictor (i.e., a good range in the number of highly effective teachers) across students, such a design should facilitate a reasonably sensitive test of the relationship between teacher characteristics and school outcomes, providing consistent longitudinal outcome data are available on the students (e.g., Project Talent data; Flanagan & Cooley, 1966). Madaus et al. (1980) discuss some of the drawbacks to longitudinal quasi-experimentation, most having to do with the inability to assign students to groups (see also Marco, Murphy, & Quirk, 1976). The fact that the researcher cannot select the criterion measures could also seriously affect the study's sensitivity to true differences.

Returning to the "ideal" design for a moment, perhaps a socially acceptable design that would be more sensitive than the quasi-design presented above is one in which students are randomly selected to all "highly effective" teachers with comparisons made between this group and those students who had exposure to less than all "highly effective" teachers. This type of quasi-experiment approximates the normal class assignment situation and thus would not disrupt the educational process. Also, and given the points

raised above about the criteria, a critical consideration is the fact that such a design would again allow the experimenter to select the most construct valid outcome measures for study in the longitudinal design.

The typical study of school effectiveness divides the criterion variance into within and between school variance. Main effects in this type of analysis for characteristics of teachers or principals are evident when there are greater differences across schools or school districts than within such aggregations. With regard to teacher effects, there is an implicit assumption that there are factors at the school, district, or state level that somehow distinguish teaching quality. For example, might some school districts have more valid selection systems or lower selection ratios (i.e., a lower ratio of applicants to openings)? Might some school districts spend a relatively greater amount of money training teachers to be more effective? Also, might teachers of potentially higher quality select certain school districts over others? Similar questions can be asked of principal or administrator characteristics. Based on our review of the literature, personnel selection and training in education is generally uniform across schools, districts and other levels of aggregation. In other words, there is little evidence that schools or districts differ in the extent to which they effectively train their personnel. Several studies have found that teachers of potentially higher quality (based on acievement test scores, for example) may select certain schools and districts over others. This latter finding, however, has little to do with school characteristics other than perhaps their physical proximity (e.g., inner-city versus suburban). Thus, in terms of personnel selection and training, there is little evidence to suggest

that between-school variance will be greater than within-school variance other than that which is related to background characteristics of teachers. The higher level of data aggregation (i.e., to the school level) for teacher and administrative characteristics thus uses a mean value to represent such characteristics when there is a great likelihood that the majority of students from any one school or district do not receive the "average" level of these characteristics. Madaus et al. (1980) state that teacher characteristics have been aggregated to the school level in "virtually every study of differential school effectiveness..." (1980, p. 90). The poor predictability in these studies for teacher and other school-related effects (e.g., resources and facilities) is certainly explainable given the level of data aggregation consistently used in such studies.

We believe the conclusions of Coleman et al. (1966), Jencks et al. (1972), Levine and Bane (1975) and others regarding the effects of schooling are unwarranted given the designs and methods employed to study effects. In our review of the literature, we failed to find an experimental or quasi-experimental design that was adequately sensitive to be able to partial out the effects of intuitively appealing, school related variables.

Applications of models of organizational effectiveness should only be empirically tested once the issues of data aggregation, the construct validity of criteria for school outcomes, and adequate sample sizes are considered thoroughly. Before these critical issues are addressed, however, of paramount concern for researchers on school effectiveness are the objectives or purposes for schooling. Until there is understanding on this most complicated of issues, researchers are in a position succintly captured by the

following passage:

> "More than any other time in our history, mankind faces a crossroads. One path leads to despair and utter hopelessness. The other, to total extinction. Let us pray that we have the wisdom to choose correctly (Allen, 1980, p. 57).

References

Abt, C. Applying cost-benefit paradigms to social program evaluations. Paper presented at the meeting of the Evaluation Research Society, Washington, D.C., 1977.

Airasian, P. W. A perspective on the uses and misues of standardized achievement tests. Measurement in Education, Fall 1979, 10, 3.

Airasian, P. W., Kellaghan, T. & Madaus, G. F. Concepts of School Effectiveness. (Contract No. NE-6-00-3-0214) National Institute of Education, Testing, Assessment and Evaluation Division, 1979.

Airasian, P. & Madaus, G. A Study of the Sensitivity of School and Program Effectiveness Measures. Report submitted to the Carnegie Corporation, New York. Chestnut Hill, Mass: Boston College, School of Education, 1976.

Albanese, R. Management: Toward Accountability for Performance. Richard D. Irwin, Inc., Homewood, Il., 1975.

Albrecht, K. Successful Management by Objectives An Action Manual. Englewood Cliffs: Prentice Hall, 1978.

Alexander, K. & Griffin, L. School district effects on academic achievement: A reconsideration. American Sociological Review, 1975, 40, 144-152.

Alkin, M. Accountability defined. In J. Roueche and B. Herrscher (Ed.) Toward Instructional Accountability. Palo Alto: Westinghouse Learning, 1973.

Allen, W. Side Effects. New York: Random House, 1980.

Altergott, B. Management-by-Objectives for the Public Schools. Unpublished doctoral dissertation, University of Indiana, 1970.

American Psychological Association, AERA, NCME. Standards for educational and psychological tests. Washington: American Psychological Association, 1974.

American Psychological Association, Division of Industrial-Organizational Psychology. Principles for the validation and use of personnel selection procedures (Second Edition). Berkeley, CA: author, 1980.

Anderson, B. & Tissier, R. Social Class School Bureaucratization and Educational Aspirations. Educational Administration Quarterly, 1973, 9, 34-49.

Anderson, G. & Nilsson, S. Studies in the reliability and validity of the critical incident technique. Journal of Applied Psychology, 1964, 48, 398-403.

Anderson, H. A Study of Certain Criteria of Teaching Effectiveness. Journal of Experimental Education, 1954, 23, 41-71.

Andes, J. Alternative Models for Urban District Organization. Educational Administration Quarterly, 1971, 1, 64-86.

Aplin, J. C. & Schoderbek, P. P. How to measure MBO. Public Personnel Management, 1976, 5, 88-95.

Austin, G. Exemplary schools and the search for effectiveness. Educational Leadership, 1979, 37, 10-14.

Averch, H., Carroll, S., Donaldson, T., Kiesling, H., and Pincus, J. How Effective is Schooling? A Critical Review and Synthesis of Research Findings. Santa Monica, California: Rand Corporation, 1972.

Barr, A., Torgerson, T., Johnson, C., Lyon, V., & Walvoord, A. The Validity of Certain Instruments Employed in the Measurement of Teaching Ability. The Measurement of Teaching Efficiency, H. Walker, ed., New York, MacMillan Co., 1935, 73-141.

Barro, S. An approach to developing accountability measures for the public schools. Phi Delta Kappan, 1970, 52, 196-205.

Basset, T. R. Its side-effects of education that count. Phi Delta Kappa, 1972, 54, 16-17.

Beck, A. & Hillman, E. Making MBO/R Work. Addison-Wesley Publishing Company, 1976.

Beckhard, R. Organizational Development. Reading, MA: Addison Wesley, 1969.

Bennis, W. G. Organization development: Its nature, origins, and prospects. Reading, Mass: Addison-Wesley, 1969.

Bennis, W. Toward a Truly Scientific Management: The Concept of Organizational Health, in Changing Organizations. New York: McGraw Hill, 1966, 32-63.

Bernardin, H. J. & Smith, P. C. A clarification of some issues regarding the development and use of behaviorally anchored rating scales. Journal of Applied Psychology, 1981, 66.

Bernhardt, R. G. A study of the relationships between teachers attitudes toward militancy and their perceptions of selected organizational characteristics of their schools. (Doctoral dissertation, Syracuse University) Ann Arbor: University Microfilms, 1972, No. 72-11, 825.

Beshers, J. Models of the educational process: A sociologist's perspective. In F. Mosteller and D. Moynihan On Equality of Educational Opportunity. New York: Random House, 1972.

Bidwell, C. The school as a formal organization. In J. March, Handbook of Organizations. Chicago: Rand McNally, 1965.

Bidwell, C. & Kasarda, J. Reply to Hannan, Freeman and Meyer, and Alexander and Griffin. American Sociological Review, 1975, 40, 152-160.

Bishop, L. & George, J. Organizational Structure: A Factor Analysis of Structural Characteristics of Public Elementary and Secondary Schools, Educational Administration Quarterly, 1975, 3, 66-80.

Blaschke, C. Performance Contracting: Who Profits Most? Bloomington, Indiana: The Phi Delta Kappa Educational Foundation, 1972.

Blau, P. & Scott, W. Formal Organization, New York: Chandler, 1962.

Bloom, B. S. Human characteristics and school learning. New York: McGraw-Hill, 1976.

Bowen, H. R. & Douglass, G. K. Efficiency in liberal education. A study of comparative instructional costs for different ways of organizing teaching-learning in a liberal arts college. New York: McGraw-Hill, 1971.

Brookover, W. B., Schweitzer, J. H., Schneider, J. M., Beady, C. H., Flood, P. K., & Wisenbaker, J. M. Elementary school social climate and school achievement. American Educational Research Journal, 1978, 15, 301-318.

Brophy, J. Stability of Teacher Effectiveness. American Educational Research Journal, 1973, 10(3), 245-252.

Bucalo, J. Personnel director . . . what you should know before recommending MBO. Personnel Journal, 1977, 56(4), 176-178.

Burns, T. Management in action. Operational Research Quarterly, 1957, 8, 45-60.

Byrnes, J. L. A study of certain relationships among perceived supervisory style, participativeness, and teacher job satisfaction. (Doctoral dissertation, Syracuse University) Ann Arbor: University Microfilms, 1973, No. 73-7790.

Calsyn, R., Fornatzky, L. & Dittmar, S. Incomplete Adoption of an Innovation: The Case of Goal Attainment Scaling. Evaluation Quarterly, 1977, 4, 127-130.

Cameron, K. Measuring organizational effectiveness in institutions of higher education. Administrative Science Quarterly, 1978, 23, 604-629.

Campbell, J. Contributions Research Can Make in Understanding Organizational Effectiveness. In Organizational Effectiveness: Theory, Research, Utilization. Spray, S. (Ed.). Kent State University Press, 1976.

Campbell, J., Bownas, D., Peterson, N., & Dunnette, M. The Measurement of Organizational Effectiveness: A Review of Relevant Research and Opinion (Contract No. N00022-73-C-0023). San Diego, Calif.: Navy Personnel Research and Development Center, 1974.

Campbell, J., Dunnette, M., Lawler, E. & Weick, E. Managerial Behavior, Performance, and Effectiveness. New York: McGraw Hill, 1970.

Carlson, S. Executive Behavior: A Study of the Work Load and the Working Methods of Managing Directors. Stockholm: Stromberg, 1951.

Carpenter-Huffman, P., Hall, G. & Summer, G. Change in Education, Cambridge, Mass: Ballinger Publishing Company, 1974.

Carr, R. W. A study of the job satisfaction of high school principals. (Doctoral dissertation, The University of Michigan) Ann Arbor: University Microfilms, 1971, No. 71-23, 719.

Chase, C. Achievement Tests. In Measurement For Educational Evaluation, Mass: Addison-Wesley Publishing Co., 1974, 188-209.

Chase, Jr., S. (Ed.) Problems in Public Expenditure Analysis. Washington, D.C.: Brookings Institution, 1968.

Cicirelli, V. G. The relevance of the regression artifact problem to the Westinghouse-Ohio Evaluation of Head Start: Reply to Campbell and Erlebacher. In J. Hellmuth (Ed.), Disadvantaged Child, Vol 3, New York: Brunner/Mayel, 1970.

Cicirelli, V. G., Evans, J. W. & Schiller, J. S. The impact of Head Start: A reply to the report analysis. Harvard Educational Review, 1970, 40, 105-129.

Cicourel, A. (Chairman). Organizational Processes in Education: Field Research on Interactions Within Educational Organizations. Report from a Conference, La Jolla, Calif., 1975.

Coleman, J. Methods and results in the IEA studies of effects of school on learning. Review of Educational Research, 1975, 45, 335-386.

Coleman, J. The evaluation of Equality of Educational Opportunity. In Mosteller and Moynihan's On Equality of Educational Opportunity. New York: Random House, 1972.

Coleman, J., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, P., & York, R. Equality of Educational Opportunity. Washington, D.C.: U.S. Government Printing Office, 1966.

Coleman, J. S. & Karweit, N. L. Information systems and performance measures in schools. Englewood Cliffs, NJ: Educational Technology Publications, 1972.

Connally, T., Conlon, E., & Deutsch, S. Organizational Effectiveness: A Multiple Constituency Approach. The Academy of Management Review, 1980, 5, 211-218.

Cooley, W. & Lohnes, P. Evaluation Research in Education. New York: Wiley, 1976.

Corwin, R. Militant Professionalism: A Study of Organizational Conflict in High Schools. New York: Appleton-Century Crofts, 1970.

Craft, J. & Birnberg, J. Human Resource Accounting: Perspective and Prospects. Industrial Relations, 1976.

Crittenden, K., Norr, J. Some remarks on "student ratings": The validity problem. American Educational Research Journal, 1975, 12(4), 429-433.

Cronbach, L. Research on Classrooms and Schools: Formulation of Questions, Design, and Analysis. Standard Evaluation Consortium, Stanford, Calif., 1976.

Cronbach, L. J. Test validation. In R. L. Thorndike (Ed.), Educational measurement (2nd ed.). Washington, D.C.: American Council on Education, 1971.

Cronbach, L. J. Course improvement through evaluation. Teacher's College Record, 1963, 64, 672-683.

Cullers, B. Unpublished study, 1970. In R. Likert, The profile of a school. Ann Arbor, Mich.: Rensis Likert Associates, 1978.

Cullers, B., Hughes, C. & McGreal, T. Administrative behavior and student dissatisfaction: A possible relationship. Peabody Journal of Education, 1973, 155-163.

Cyert, R. M. & March, J. G. A Behavioral Theory of the Firm. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1963.

Cunningham, J. Approaches to the evaluation of organizational effectiveness. Academy of Management Review, 1977, 2, 463-474.

Cytrynbaum, S., Ginath, Y., Birdwell, J., & Brandt, L. Goal attainment scaling: A critical review. Evaluation Quarterly, 1979, 3, 5-40.

Davis, S. & Gross-Davis, B. Assessing teacher effectiveness through student learning gains. California Journal of Educational Research, 1974, 25, 186-191.

Donlan, H. K. The leader-management system of the elementary school principal and its relation to student achievement. (Doctoral dissertation, New York University) Ann Arbor: University Microfilms, 1978, No. 79-18, 840.

Dornbusch, S. & Scott, W. Evaluation and the Exercise of Authority: Theory of Control Applied to Diverse Organizations. San Francisco: Jossey-Bass, 1975.

Drucker, P. The Practice of Management. New York: Harper-Row, 1954.

Dunnette, M. D. & Borman, W. C. Personnel sel-ction and classification systems. Annual Review of Psychology, 1979, 30, 477-525.

Dyer, H. Toward objective criteria of professional accountability in the schools of New York City. Phi Delta Kappan, 1970, 52, 206-211.

Eads, Jr., A. A Study of the Attitudes of Teachers Toward a "Supervision by Objectives" Teacher Evaluation. Unpublished doctoral dissertation, University of South Carolina, 1974.

Ebel, R. Standardized Achievement Tests: Uses and Limitations. In Readings in Psychological and Educational Testing, L. Aiken, ed., Boston: Allyn and Bacon, 1973, 17-22.

Educational Research Service, Inc. Teacher Performance Contracts, Arlington, Virginia: Educational Research Service, 1974.

Educational Testing Service, Proceedings of the Conference on Educational Accountability, Washington, D.C.: 1971.

England, G. Personal value systems of American managers. Academy of Management Journal, 1967, 10, 53-68.

England, G. Organizational Goals and Expected Behavior of American Managers. Minneapolis: University of Minnesota, Industrial Relations Center. Unpublished report, 1966.

Equal Employment Opportunity Commission. Guidelines on employee selection procedures. Federal Register, 1970, 35: 12333-12336.

ERIC. ERIC abstracts on educational planning. Available through ERIC Clearinghouse, 1973.

Etzioni, A. A Comparative Analysis of Complex Organizations. New York: The Free Press, 1975.

Etzioni, A. Modern Organizations. Englewood Cliffs, NJ: Prentice-Hall, 1964.

Etzioni, A. Two approaches to organizational analysis: A critique and a suggestion. Administrative Science Quarterly, 1969, 5, 257-278.

Fanshel, S. & Bush, J. A health-status index and its application to health-services outcomes. Operations Research, 1970, 18, 1021-1066.

Farkas, G. Specification, residuals, and contextual effects. Sociological Methods and Research, 1974, 2, 333-363.

Fincher, C. Planning models and paradigms in higher education. Journal of Higher Education, 1972, 43, 754-767.

Finn, J. D. A general model for multivariate analysis. New York: Holt, Rinehart, & Winston, 1974.

Firebaugh, G. A rule for inferring individual-level relationships from aggregate data. American Sociological Review, 1978, 43, 557-572.

Flamholtz, E. Human Resource Accounting: A Review of Theory and Research. Journal of Management Studies, 1974.

Flanagan, J. The critical incident technique. Psychological Bulletin, 1954, 51, 327-358.

Flanagan, J. C. & Colley, W. W. Project talent: One-year follow-up studies. Pittsburgh, PA: School of Education, University of Pittsburgh, 1966.

Fobert, R. The voucher plan: Is it worth the investment? The School Administrator, 1977.

Foster, J. P. Perceptions of Principals' Behavior as Rated by Teachers, Students, and Principals in the Junior High School in Chattanooga, Tennessee, 1976, No. TSZ 76-23, 216.

Freeman, J. H. The unit of analysis in organizational research. In M. W. Meyer (Ed.), Environments and organizations. San Francisco: Jossey-Bass, 1978.

Freeman, M. and Roney, A. Unpublished study, 1980. In R. Likert, The profile of a school. Ann Arbor, Mich.: Rensis Likert Associates.

Fremont, J. Using Parsons' functional analysis in the study of public organizations. Administrative Science Quarterly, 1975, 20, 59-70.

French, W. Organization development: Objectives, assumptions, and strategies. In N. Margulies and A. Raia (Eds.), Organizational Development: Values, Processes, and Technology. New York: McGraw Hill, 1972.

Frey, P., Leonard, D. & Beatty, W. Student ratings of instructions: validation research. American Educational Research Journal, 1975, 12, 435-444.

Garfinkel, I. & Gramlich, E. A Statistical Analysis of the OEO Experiment in Educational Performance Contracting. Journal of Human Resources, 1973, 8, 275-305.

Garms, W. A benefit-cost analysis of the upward bound program. Journal of Human Resources, 1971, 6, 206-220.

Georgopoulos, B. & Tannenbaum, A. A study of organizational effectiveness. American Sociological Review, 1957, 22, 534-540.

Ghiselli, E. Dimensional problems of criteria. Journal of Psychology, 1956, 40, 1-4.

Ghorpade, J. Assessment of Organizational Effectiveness. Pacific Palisades, Calif.: Goodyear Publishing Company, Inc., 1971.

Gibson, A. K. The achievement of sixth grade students in a mid-western city. (Doctoral dissertation, The University of Michigan) Ann Arbor: University Microfilms, 1974, No. 74-15, 729.

Glaser, R. & Nitko, A. J. Measurement in learning and instruction. In R. L. Thorndike (Ed.), Educational Measurement (2nd Ed.). Washington, D.C.: American Council on Education, 1971.

Glennan, T. Evaluating federal manpower programs: Notes and observations. In P. Rossi and W. Williams (Eds.), Evaluating Social Programs, New York: Seminar, 1972.

Glick, W. H. & Roberts, K. H. Hypothesized interdependence, assumed independence. Paper presented at the Annual Meeting of the American Psychological Association, San Francisco, 1978.

Gramlich, E. & Koshel, P. Educational Performance Contracting, Washington, D.C.: The Brookings Institution, 1975.

Gray, F. & Burns, M. Does "Management-by-Objectives" Work in Education? Educational Leadership, 1979, 414-417.

Grayson, L. Costs, benefits, effectiveness: Challenge to educational technology. Science, 1972, 175, 1216-1222.

Gross, N. & Herriott, R. Staff Leadership in Public Schools. New York: Harper, 1965.

Guba, E. & Bidwell, C. Administrative relationships: Teacher effectiveness, teacher satisfaction, and administrative behaviors: A study of the school as a social institution. Studies in Educational Administration, Chicago: University of Chicago, 1957.

Guion, R. The employee load of first line supervisors. Personnel Psychology, 1953, 7, 223-244.

Hall, H. Professional preparation and teacher effectiveness. The Journal of Teacher Education, 1964, 15, 72-76.

Hannan, M. Aggregation and Disaggregation in Sociology. Lexington, Mass.: Heath, 1971.

Hannan, M. T., & Burstein, L. Estimation from grouped observations. American Sociological Review, 1974, 39, 374-392.

Hannan, M., Freeman, J. & Meyer, J. Specification of models for organizational effectiveness. American Sociological Review, 1976, 41, 136-143.

Hanson, E. The Modern Educational Bureaucracy and the Process of Change, Educational Administration Quarterly, 1975, 3, 21-36.

Hanushek, E. A Reader's Guide to Educational Production Functions. Available through NIE, 1977.

Harris, C. W. (Ed.) Problems In Measuring Change. Madison: University of Wisconsin Press, 1963.

Hartmark, L. Accountability, Efficiency, and Effectiveness in the State University of New York, SUNY-Albany: Comparative Development Studies Center, 1975.

Hatry, H. Measuring the effectiveness of nondefense public programs. Operations Research, 1970, 18, 772-784.

Haynes, P. D. A comparison of perceived organizational characteristics between selected work stoppage and non-work stoppage school districts in the state of Michigan. (Doctoral dissertation, Western Michigan University) Ann Arbor: University Microfilms, 1972, No. 72-14, 182.

Heil, L. & Washburne, C. Brooklyn College Research in Teacher Effectiveness, The Journal of Educational Research, 1962, 55(8), 347-350.

Herman, J. How and why a school district implemented MBO. NASSP Bulletin, 1978, 62, 82-86.

Hersey, P. NASSP's assessment center-from concept to practice. The Industrial-Psychologist, 1977, 15(1), 22-23.

Hilfiker, L. The Relationship of School System Innovativeness to Selected Dimension of Interpersonal Behavior in Eight School Systems. Madison: Research and Development Center, 1969.

Horne, J. & Lupton, T. The work activities of "middle" managers. Journal of Management Studies, 1965, 1, 14-33.

Horsman, J. Unpublished study, 1973. In R. Likert, The profile of a school. Ann Arbor. Mich.: Rensis Likert Associates, 1978.

Hottleman, G. Performance contracting is a hoax. The Massachusetts Teachers, 1971, 4-10.

Hoy, W., Newland, W. & Blazovsky, R. Subordinate Loyalty to Superior, Esprit, and Aspects of Bureaucratic Structure, Educational Administration Quarterly, 1977, 13, 71-85.

Hoy, W., Tarter, C. & Forsyth, P. Administrative Behavior and Subordinate Loyalty: An Empirical Assessment, Journal of Educational Administration, 1978, 16, 29-38.

Humble, J. How to Manage by Objectives, New York: AMACOM, 1973.

Ingenkamp, K. Educational Assessment. Windsor, Berks: NFER Publishing Company, 1977.

Jackson, J. & Mathis, R. Management-by-Objectives: Promises, Pitfalls, and Possibilities, Personnel Administration/Public Personnel Review, 1972, 72-75.

James, L. R. Criterion models and construct validity for criteria. Psychological Bulletin, 1973, 80, 75-83.

Jencks, C., Smith, M., Acland, H., Bane, M., Cohen, D., Gintis, H., Heyns, B., & Michesson, S. Inequality: A Ressessment of the Effects of Family and Schooling in America. New York: Basic Books, 1972.

Jones, N. The relative effectiveness of MBO trained counselors. Unpublished manuscript, 1977.

Katz, D. & Kahn, R. The Social Psychology of Organizations. New York: Wiley, 1978.

Katz, D. & Kahn, R. The Social Psychology of Organizations. New York: John Wiley & Sons, 1966.

Kay, B. R. Key factors in effective foreman behavior. Personnel, 1959, 36, 25-31.

Kellaghan, T. Measuring school effectiveness. In R. Summer (Ed.), Monitoring National Standards of Attainment in Schools. Slough, England: NFER Publishing Co., 1977.

Kelly, J. The study of executive behavior by activity sampling. Human Relations, 1964, 17, 277-287.

Key, R. C. A study of perceived organizational characteristics in persistent disagreement school districts and nonpersistent disagreement districts in the San Francisco Bay Area. (Doctoral dissertation, University of Southern California) Ann Arbor: University Microfilms, 1974, No. 74-21, 483.

Kirchhoff, B. A. Organizational effectiveness measurement and policy research. Academy of Management Review, July, 1977, 347-355.

Klein, S. P. The uses and limitations of standardized tests in meeting the demands for accountability. UCLA Evaluation Comment, Center for the Study of Evaluation, 2(4), 1971.

Klitgaar, R. E. Going beyond mean in educational-evaluation. Public Policy, 1975, 23, 59-79.

Kraemer, H. C. Individual and ecological correlation in a general context. Behavioral Science, 1978, 23, 67-72.

Kiresuk, I. & Sherman, R. Goal attainment scaling: A general method for evaluating community mental health programs. Community Mental Health, 1968, 4, 443-453.

Ladouceur, J. School management profile and capacity for change. (Unpublished doctoral dissertation, University of Toronto, 1973).

Landy, F. & Farr, J. Performance rating. Psychological Bulletin, 1980, 87, 72-107.

Langland, Unpublished study, 1978. In R. Likert, The profile of a school. Ann Arbor, Mich.: Rensis Likert Associates, 1978.

Lesher, H. Is MBO appropriate in the academic setting? Educational Leadership, 1979, March, 414-418.

Latham, G. P. & Yukl, G. A. A review of research on the application of goal setting in organizations. Academy of Management Journal, 1975, 18(4), 824-845.

Law, L. Educational Production Functions. Available from NIE, 1977.

Lawrence, P. R. & Lorsch, J. W. Organization and environment. Homewood, Il.: Richard D. Irvin, 1969.

Lepkowski, Sr., M. L. Cooperative decision making as related to supportive relations and communication in the senior high school. (Doctoral dissertation, University of Buffalo) Ann Arbor: University Microfilms, 1970, No. 70-22, 106.

Levin, H. Concepts of economic efficiency and educational production. In Froomkin, J. et al.'s Education As An Industry. Cambridge, MA: Ballinger Publishing Co., 1976.

Levin, H. Cost-effective analysis in evaluation research. In M. Guttentag and E. Struenissy (eds.). Handbook of Evaluation Research, Beverly Hills: Sage, 1975, 2.

Levine, D. M. & Bane, M. J. Introduction. In D. M. Levine & M. J. Bane (Eds.), The "inequality" controversy: Schooling and distributive justice. New York: Basic Books, 1975.

Levinson, H. Management by whose objectives? Harvard Business Review, 1970, 48, 125-134.

Lewy, A. Discrimination among individuals vs. discrimination among groups. Journal of Educational Measurement, 1973, 10, 19-24.

Lieberman, M. An overview of accountability. Phi Delta Kappan, 1970, 52, 194-195.

Likert, R. The Likert Profile of a School, 2nd ed., Ann Arbor, Mich.: Rensis Likert Associates, 1978.

Likert, R. Human Resource Accounting: Measuring Positional Replacement Costs. Human Resource Management, 1973.

Likert, R. The Human Organization: Its Management and Value. New York: McGraw-Hill, 1967.

Likert, R. New Patterns of Management. New York: McGraw-Hill, 1961.

Lopez, F. Accountability in education. Phi Delta Kappan, 1970, 52, 231-235.

Lortie, D. The balance of control and autonomy in elementary school teaching. In A. Etzioni (Ed.), The Semi-professions and Their Organization. New York: Free Press, 1969.

Luthan, F. How to apply MBO. Public Personnel Management, 1976.

MacKillican, Unpublished study, 1975. In R. Likert, The profile of a school. Ann Arbor, Mich.: Rensis Likert Associates, 1978.

Macy, B. & Mirvis, P. A methodology for assessment of quality of work life and organizational effectiveness in behavioral-economic terms. Administrative Science Quarterly, 1976, 21, 212-225.

Madaus, G. F., Airasian, P. W. & Kellaghan, T. School effectiveness: A reassessment of the evidence. New York: McGraw-Hill Book Co., 1980.

Madaus, G., Kellaghan, T., Rakow, E. & King, D. The sensitivity of measures of school effectiveness. Harvard Educational Review, 1979, 49, 207-228.

Mahoney, T. & Weitzel, W. Managerial Models of Organizational Effectiveness. Administrative Science Quarterly, 1969, 14, 357-365.

Malsam, M. MBO: One district's success story. American Schools and Universities, 1979, 34-38.

Mangum, G. L. Manpower policies for a democratic-society-national manpower council. Book Review, Journal of Human Resources, 1967, 21, 265-267.

Marco, G. L., Murphy, R. T. & Quirk, T. A classification of methods of using student data to assess school effectiveness. Journal of Educational Measurement, 1976, 13, 243-252.

Marsh, R. Management-by-objectives: A multifaceted faculty evaluation model. Educational Technology, 1979, 19, 44-48.

Mayrhofer, A. Performance contracting for Instruction. In Science and Jantz, Accountability in American Education. Boston: Allyn-Bacon, 1972.

McClelland, D. Testing for competence rather than for "intelligence". American Psychologist, 1973, 28, 1-14.

McDonald, F. A Design For An Accountability System For The New York City School System. Educational Testing Service, Princeton, New Jersey, 1972.

McLagan, P. Competency models. Training and Development Journal, 1980, 34, 22-26.

McNeil, J. & Popham, W. The assessment of teacher competence. In W. Travers (Ed.) Second Handbook of Research on Teaching. Rand McNally: 1973.

Mecklenburger, J. & Wilson, J. The performance contract in Gary. Phi Delta Kappan, 1971, 52, 406-411.

Meeth, R. Quality Education for Less Money. San Francisco: Jossey-Bass, 1974.

Mehrens, W. A. & Ebel, R. L. Some comments on criterion-referenced and norm-referenced achievement tests. NCME Measurement in Education, 1979, 10, No. 1.

Metfessel, N. & Michael, W. A paradigm for developing valid measurable objectives in the evaluation of educational programs in colleges and universities. Educational and Psychological Measurement, 1967, 27, 373-383.

Miller, D. R. Unpublished manuscript, 1970. In R. Likert, The profile of a school. Ann Arbor, Mich.: Rensis Likert Associates, Inc., 1978.

Mirvis, P. & Lawler, E. Measuring the financial impact of employee attitudes. Journal of Applied Psychology, 1977, 62, 1-8.

Miskel, C., Fevurly, S. & Stuart, R. Organizational structures and processes, perceived school effectiveness, loyalty, and job satisfaction. Educational Administration Quarterly, 1979, 15, 97-118.

Miskel, C. & Gerhardt, E. Perceived Bureaucracy and Teacher Conflict: Central Life Interests, Voluntarism, and Job Satisfaction. Journal of Educational Administration, 1974, 10, 84-97.

Mohr, L. (Chairman). Administrative Structure, Effectiveness, and Efficiency. A Prospectus for Research in Organizational Aspects of Education prepared for the National Institute of Education, 1975.

Mohr, L. The concept of organizational goal. American Political Science Review, 1973, 77(2), 470-481.

Morall, H. H. The relationship between perceived participation in school management and morale of selected black and nonblack teachers and students in Volusia county, Florida senior high schools. (Doctoral dissertation, University of Miami) Ann Arbor: University Microfilms, 1974, No. 74-23, 405.

Morgenstern, O. On the Accuracy of Economic Observations. Princeton, NJ: Princeton University Press, 1963.

Mott, P. The Characteristics of Effective Organizations. New York: Harper & Row, 1972.

Murphy, M., Bishop, L., George, J. Defining organizational properties of schools: A focus on structure. Paper presented at the Annual Meeting of the American Educational Research Association, Washington, D.C., 1975.

Nash, R. Commitment to competency: The new fetishism in teacher education. Phi Delta Kappan, 1970, 52, 240-243.

Nation's Schools. Performance contracting: Clouds and controversy over Texarkana. Nation's Schools, 1970, 86, 85-88.

Naumann-Etienne, M. Bringing about open education: Strategies for innovation and implementation. (Doctoral dissertation, The University of Michigan) Ann Arbor: University Microfilms, 1975, No. 75-20, 417.

Nirenberg, J. A Comparison of the Management Systems of Traditional and Alternative Public High Schools, Educational Administration Quarterly, 1977, 13, 86-104.

Nunnally, J. Psychometric Theory. New York: McGraw-Hill, 1978.

Odetola, T. Organizational Structure and Student Alienation. Educational Administration Quarterly, 1972, 8, 15-26.

Odiorne, G. Management-By-Objectives. San Francisco: Pitman Publishing Corp., 1965.

O'Donoghue, M. Economic dimensions in education. Dublin: Gill and Macmillan, 1971.

O'Neill, H. & Kubany, A. Observation methodology of supervisory behavior. Personnel Psychology, 1959, 12, 85-96.

O'Neill, J. Resource Use in Higher Education: Trends in Outputs and Inputs. New York: McGraw-Hill, 1971.

Parsons, T. Structures and process in modern societies. Glencoe, IL: The Free Press, 1960.

Parsons, T. Suggestions for a sociological approach to the theory of organizations. Administrative Science Quarterly, 1956, 1, 63-85.

Paul, S. An application of cost-benefit analysis to management education, in W. Niskanen et al., Benefit-Cost and Policy Analysis. Chicago: Aldine, 1973, 374-392.

Pellegrin, R. Schools as work settings. In R. Dubin Handbook of Work, Organization, and Society. Chicago: Rand-McNally, 1976, 343-376.

Pellegrin, R. Some organizational characteristics of multi-unit schools. Working Paper No. 22. Madison: Wisconsin Research and Development Center, University of Wisconsin, 1968.

Pellegrin, R. The place of research in planned change. In R. Carlson et al. Change Processes In the Public Schools. Eugene, Oregon: Center for the Advanced Study of Educational Administration, Univ. of Oregon, 1965.

Pennings, J. Dimensions of organizational influence and their effectiveness correlates. Administrative Science Quarterly, 1976, 21, 688-699.

Pennings, J. The relevance of the Structural-Contingency Model for organizational effectiveness. Administrative Science Quarterly, 1975, 20, 393-410.

Perrow, C. Complex Organizations: A Critical Essay. Glenview, IL: Scott, Foresoman, 1972.

Perrow, C. Goals in complex organizations. American Sociological Review, 1961, 6, 854-865.

Pfeffer, J. Usefulness of the concept. In P. S. Goodman and J. M. Pennings (Eds.), New perspectives on organizational effectiveness. San Francisco: Jossey-Bass, 1977.

Popham, W. Educational Evaluation. Englewood Cliffs, NJ: Prentice-Hall, 1975.

Popham, W. J. Criterion-Referenced Instruction. California: Fearon Publishers, 1970.

Posavac, E. & Carey, R. Program Evaluation: Methods and Case Studies. Englewood Cliffs: Prentice-Hall, 1980.

Price, J. The study of organizational effectiveness. Sociological Quarterly, 1972, 13, 3-15.

Radnor, M. (Chairman). Administration and management in educational organizations. Discussion of Issues and Proposal for a Research Program for the National Institute of Education prepared at a Workshop, Watsonville, Calif., 1974.

Ratsoy, E. Participative and Hierarchical Management of Schools: Some Emerging Generalizations. Journal of Educational Administration, 1973, 11(2), 161-170.

Reimann, B. Organizational effectiveness and management's public values: A canonical analysis. Academy of Management Review, 1975, 18(2), 224-241.

Rice, A. The Enterprise and Its Environment. London: Tavistock, 1963.

Riedel, J. E. A comparison of principal, teacher and student perceptions of selected elementary school principals' effectiveness. (Doctoral dissertation, University of Southern California) Ann Arbor: University Microfilms, 1974, No. 74-21, 504.

Rivlin, A. Systematic Thinking For Social Action. Washington, D.C.: The Brookings Institution, 1971.

Roberts, K. H., Hulin, C. L., & Rousseau, D. M. Developing an interdisciplinary science of organizations. San Francisco: Jossey-Bass, 1978.

Robinson, J. The observation report-a help or a nuisance? NASSP Bulletin, 1978, 62(416), 22-26.

Robinson, W. Ecological correlations and the behavior of individuals. American Sociological Review, 1950, 15, 351-357.

Rosenshine, B. The Stability of Teacher Effects Upon Student Achievement, Review of Educational Research, 1970, 40(5), 647-662.

Rosenshine, B. & Furst, J. Current and future research on teacher performance criteria. In B. W. Smith (Ed.) Research on Teacher Education: A Symposium. Englewood Cliffs: Prentice-Hall, 1971.

Rossano, M. A Study of Management by Objectives as Perceived and Practiced by Junior College Administrators. Unpublished doctoral dissertation. University of Texas at Austin, 1975.

Rossi, P., Freeman, H., & Wright, S. Evaluation: A Systematic Approach. Beverly Hills: Sage Publications, 1979.

Schein, E. Organizational Psychology. Englewood Cliffs, NJ: Prentice-Hall, 1965.

Schmidt, F. L., Hunter, J. E., & Urry, V. W. Statistical power in criterion-related validity studies. Journal of Applied Psychology, 1976, 61, 473-485.

Schneider, B. Work climates: An interactionist perspective. Paper presented at the Third Annual Symposium on Applied Behavioral Science, Virginia Polytechnic Institute and State University, May, 1980.

Schneider, B. Different levels of analysis in data aggregation: Problems, pitfalls, and a potential solution for survey research. Paper presented at the Annual Meeting of the American Psychological Association, 1978.

Scriven, M. The methodology of evaluation. In R. Tyler, R. Gagne, and M. Scriven (eds.), Perspectives in Curriculum Evaluation. Chicago: Rand McNally, 1967.

Scott, C. Some remarks on "student ratings": Validation. American Educational Research Journal, 1975, 12(4), 444-447.

Selynick, P. TVA and the grass roots. Berkeley: University of California, 1953.

Shavelson, R. & Russo, N. Generalizability of Measures of Teacher Effectiveness, Educational Research, 1975, 19(3).

Shepard, L. Standard setting issues and methods. Applied Psychological Measurement, 1980, 4, 447-467.

Shetty, Y. & Carlisle, H. Application of management-by-objectives in a university setting: An exploratory study in faculty reactions. Educational Administration Quarterly, 1974, 10, 65-81.

Shetty, Y. & Carlisle, H. Organizational correlates of a management-by-objectives program. Academy of Management Journal, 1974, 17, 155-160.

Silberman, C. Crisis in the Classroom: The Remaking of American Education. New York: Random House, 1970.

Smallridge, R. J. A study of relationships between the perceived management system of elementary schools and the personal needs satisfaction of teachers. (Doctoral dissertation, George Peabody College for Teachers) Ann Arbor: University Microfilms, 1972, No. 72-25, 404.

Smith, J. B. An Investigation of Management-by-Objectives in Units of Selected Higher Educational Institutions. Unpublished doctoral dissertation, Rutgers University, 1975.

Smith, M. C. The relationship between the participative management style of elementary school principals as perceived by their teachers and the level of teacher morale. (Unpublished doctoral dissertation, University of Southern California, 1975).

Smith, P. & Kendall, L. Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. Journal of Applied Psychology, 1963, 47, 149-155.

Snow, R. E. Representative and quasi-representative designs for research on teaching. Review of Educational Research, 1974, 44, 265-292.

Soar, R. & Soar, R. Problems in Using Pupil Outcomes for Teacher Evaluation, in Standardized Testing Issues: Teachers' Perspective, N.E.A., Washington, D.C., 1977, 52-57.

Stanchfield, J. A place to stand in literacy or the principal can make a difference. Thrust for Educational Leadership, 1976, 6(2), 22-23; 31.

Steers, R. Organizational Effectiveness: A Behavioral View. Pacific Palisades, Calif.: Goodyear, 1977.

Steers, R. Methodological issues in evaluating organizational effectiveness. Paper presented at the annual meeting of the American Psychological Association, Washington, D.C., 1976.

Stewart, R. Managers and Their Jobs. London: Macmillan, 1967.

Stone, F. Investment in Human Resources at AT&T. Management Review, 1972, 23-27.

Stow, S. & Sweeney, J. A comprehensive three year process for planning a system of accountability can ensure discriminating and valid results. Educational Leadership, 1981, 539-541.

Tebay, J. E. Planning that begins and ends with people. Management Review, 1973, 62, 47-51.

Thompson, J. P. & McEwen, W. J. Organizational goals and environment. American Sociological Review, 1958, 23, 23-30.

Tomassini, L. Assessing the Impact of Human Resource Accounting: An Experimental Study in Managerial Decision Preferences. Accounting Review, 1977.

Vanzandt, C. The effects of MBO accountability on guidance programs in Maine. Unpublished manuscript, 1977.

Wagstoff, L. H. The relationship between administrative systems and interpersonal needs of teachers, (Doctoral dissertation, University of Oklahoma) Ann Arbor: University Microfilms, 1970, No. 70-2343.

Watson, J. Kids as raters: Can they evaluate? Instructor, 1974, 83(8), 37.

Weeks, Unpublished study, 1978. In R. Likert, The profile of a school. Ann Arbor, Mich.: Rensis Likert Associates, 1978.

Weick, K. Educational organizations as loosely coupled systems. Administrative Science Quarterly, 1976, 21, 1-19.

Werts, C. & Linn, R. Considerations when making inferences within the analysis of covariance model. Educational and Psychological Measurement, 1971, 31, 407-416.

Williams, R. A Description of Some Executive Abilities by Means of Critical Incident Technique. Unpublished doctoral dissertation, Columbia University, 1956.

Wolfe, B. A cost-effective analysis of reduction in school expenditures: An application of an educational production function. Journal of Educational Finance, 1977, 2, 407-418.

Yee, A. (Ed.) Perspectives on Management Systems Approaches in Education: A Symposium. Englewood Cliffs: Educational Tech. Publications, 1973.

Yuchtman, E. & Seashore, S. E. A system resource approach to organizational effectiveness. Administrative Science Quarterly, 1967, 32, 377-395.

Footnote

[1]The authors gratefully acknowledge the assistance of the following individuals: Bruce Etringer, Rick Fort, Dr. Jim Fortune, and Meta Lowe for contributions to various portions of this project. We would particularly like to thank Dr. M. Bruce Mackenzie-Haslam for his patience and guidance throughout this project.

Table 1

Applicability of Goal-Oriented Models to School Effectiveness

MODEL	THEORETICAL REFERENCE	FOCUS	EDUCATIONAL EXAMPLE	OUTCOMES	PROBLEMS WHEN APPLIED TO EDUCATION	FEATURES APPLICABLE TO EDUCATION
Rational-Goal	Etzioni (1964); Perrow (1961)	Effectiveness defined by achievement of prescribed goals.	Blaschke (1972); Gramlich & Koshel (1975)	Pupil gain score on: standardized norm-referenced achievement and ability tests; criterion-referenced tests; curriculum-based tests.	Neglects conflicting goals of multiple constituencies; little evidence to support performance contracting.	Formation of dominant coalitions to conciliate conflicts; goals can be set for non-cognitive factors as well.
Management-by-Objectives	Drucker (1954)	Effectiveness measured by attainment of specific goals.	Altergott (1970); Eads (1974); Shetty & Carlisle (1974)	Pupil gain score on: standaridized norm-referenced achievement and ability tests; criterion-referenced tests; curriculum-based tests.	Focuses measurement on more easily quantifiable objectives; not easily tied in with individual rewards.	Limited applications to education have been positive.
Functional	Parsons (1960)	Do organizational needs serve client groups?	Hanushek (1977)	Pupil gain score on: standardized norm-referenced achievement and ability tests; criterion-referenced tests; curriculum-based tests; ratings of behavior.	Little empirical support to justify "pattern maintenance" functions toward goal attainment, unidirectionality of influence on client groups; goal measurement.	Considers conflicting "client" groups in goal attainment.
Cost-Benefit	Rivlin (1971)	Efficiency the central variable.	Wolfe (1977); Paul (1973)	Pure cost measures-expenditure per pupil; Workload measures-number of teachers per pupil; Physical standards-students per classroom; Pupil gain scores-training, turnover, costs.	Measuring efficiency; effects on non-measured factors; composite criterion measurement.	Aid in fiscal decision-making; relative program evaluation.

Table 2

Applicability of Systems-Models to School Effectiveness

MODEL	THEORETICAL REFERENCE	FOCUS	EDUCATIONAL EXAMPLE	OUTCOMES	PROBLEMS WHEN APPLIED TO EDUCATION	FEATURES APPLICABLE TO EDUCATION
Systems-Resource	Katz & Kahn (1978); Yuchtman & Seashore (1967)	Effectiveness studied in terms of efficiency or exploitation.	Levin (1976); Klitgaard (1975); Fincher (1972)	Output/input ratio (efficiency); number of graduates; percent of unemployed graduates; ability to compete for valued resources.	Identifying an acceptable standard of efficiency.	Consideration of output measures in context of suprasystem.
Functional-Structures	Pennings (1975)	Effectiveness a function of technological structural interactions.	Mohr (1973); Grayson (1972)	Formalization; bureaucratic structure; climate; leader behavior; motivational forces.	Concern with organizational security and continuity of policy; profitability most common criterion.	Relationship considered between teacher evaluative structures and output measures.
Managerial Process	Reiman (1975); Cyert & March (1963)	Effectiveness considered in context of managerial functions, abilities, and behaviors.	Barro (1970); Dyer (1970); Yee (1973)	Teacher loyalty and morale; teacher satisfaction.	Little empirical support for relationships between managerial behavior and output variables.	Methodologies proposed to study school administrator behavior; managerial dimensions identified that may be applicable.
Organizational Development	Beckhard (1969); French (1972); Bennis (1969)	Effectiveness assumed given certain organizational characteristics (e.g., upward communication, participation in decision-making).	Radnor (1974); DiCaprio (1974)	Teacher satisfaction; reactive questionnaire.	Effectiveness not measured in terms of school output measures such as achievement of non-cognitive variables.	Relationship between "healthy" system and teacher and student output measures makes good conceptual sense.
Likert-ISR	Likert (1967)	Effectiveness assumed given high participative management style.	Austin (1979)	Teacher satisfaction; student & teacher morale; labor relations; intraorganizational communication & trust; student performance(rare).	Effectiveness typically not measured with "hard" criteria.	System 4 management proven effective in other service-oriented, non-profit organizations; also limited support in education.

Table 3

Research With the Likert-ISR Model of Effectiveness

SOURCE	SAMPLE	POS FORM	RELATED TO	RELATIONSHIP FOUND*
Gibson (1974)	Sixth grade boys	Teachers, principals	Achievement test scores	Higher achievement test scores.
Key (1974)	4 school districts	Board members; Superintendents; Administrative members; principals; teachers	Labor relations	Teachers & Staff less likely to have "persistent disagreements".
Haynes (1972)	10 school systems	Board members; Superintendents; Administrative staff; principals; teachers	Teacher strikes	Less likely to strike (teachers).
Bernhardt (1972)	979 teachers	Teacher	Teacher militancy	Less teacher militancy.
Miller (1970)	82 principals 329 teachers 1099 students	Principals, teachers, students	Motivation; Attitudes & commitment to institution; Frustration; Confidence & Trust, Communication	•Higher motivation of students & teachers. •More favorable attitudes & greater commitment. •Less frustration of teachers & students. •Greater confidence & trust among persons in the school. •Better communication in all directions.
Byrnes (1973)	Teachers in seven high schools	Teacher	Job satisfaction	Greater job satisfaction.

* The closer to System 4 as compared to System 1.

Table 3 (continued)

SOURCE	SAMPLE	POS FORM	RELATED TO	RELATIONSHIP FOUND*
Wagstaff (1970)	High school teachers & principals	Teacher & principal	Interpersonal needs of teachers	Better teacher sense of need satisfaction.
Smallridge (1972)	250 elementary teachers	Teacher	Need satisfaction	Greater need satisfaction.
Carr (1971)	94 high school principals	Principal	Satisfaction	Greater satisfaction with their supervision, co-workers, work, pay & promotion.
Cullers, Hughes & McGreal (1973)	Students in six high schools	Principal & teacher	Student satisfaction	Greater student satisfaction.
Morall (1974)	106 H.S. teachers 120 H.S. students	Teacher & students	Teacher & student morale	Higher morale of students and teachers.
Smith (1975)	17 principals 254 teachers	Principal	Teacher morale, measured by the Purdue Teacher Opinionnaire	Higher teacher morale.
Lepkowski (1970)	High school teachers	Teacher	Communication	Better communications in all directions within the school.
Riedel (1974)	Elementary school teachers & principals	Teacher	Evaluators' judgment of principals	Principals evaluated higher were closer to System 4 schools.
Naumann-Etienne (1975)	4 elementary schools	-	"Open education" & traditional schools	Open schools, closer to System 4.

Table 3 (continued)

SOURCE	SAMPLE	POS FORM	RELATED TO	RELATIONSHIP FOUND*
Ladouceur (1973)	523 teachers	Teacher	Schools capacity to change	Closer to System 4, the more capable the school was of changing.
Donlan (1978)	18 schools 6th grade students	Teacher	School Achievement Scores N.Y. State Pupil Evaluation Program (PEP)	Greater the PEP score improvement.
Langland (1978)	613 teachers	Teacher	Conflict, grievance rates	Less conflict in negotiating contracts; no relationship with grievance rates.
MacKillican (1975)	469 K-6 teachers	Teacher	Use of open education	Higher the openness of the classroom.
Weeks (1978)	1403 principals and teachers in 33 schools	Teacher	Organizational climate, measured by the Profile of Conflict Characteristics	Closer teachers saw the principal's conflict management and behavior.
Horsman (1973)	14 principals, 8 vice principals, 206 teachers, 912 students	Principals, teachers, students	Educational performance	No significant correlation between POS scores and educational performance.
Freeman, Martin & Roney (1980)	2056 teachers, principals & office staff	Teacher, principal, Central Office Staff	Satisfaction & militancy	Higher System 4 = greater satisfaction = less militancy for all groups.
Cullers (1970)	37 Superintendents 44 Secondary principals, 78 Elementary principals	Superintendent, principal	Position in hierarchy	Superintendents perceive themselves as low System 4; Principals perceive themselves as high System 3, of elementary principals closer to System 4.